

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2004, 07:50:53 ; Search time 17 Seconds
(without alignments)
2468.738 Million cell updates/sec

Title: US-09-744-314-6
Perfect score: 4231
Sequence: 1 MEEPGATPQVILGLVLELR.....GLPPSNEPATEHPPEQDET 806

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3923	92.7	804	1	MEAG6 HUMAN
2	276.5	6.5	1959	1	MYH9 CHICK
3	274.5	6.5	1130	1	YL17 CAEL
4	272	6.4	3210	1	CENF HUMAN
5	271.5	6.4	1875	1	MLP1 YEAST
6	270.5	6.4	1976	1	MYHA BOVIN
7	270	6.4	1357	1	KYNI HUMAN
8	268	6.3	1957	1	SPOF SCHPO
9	266.5	6.3	1976	1	MYHA HUMAN
10	266	6.3	1790	1	US01 YEAST
11	265	6.3	1411	1	EEA1 HUMAN
12	258.5	6.1	1976	1	MYHA RAT
13	258	6.1	2017	1	MYSN DROME
14	257	6.1	1583	1	GCC2 HUMAN
15	254	6.0	1679	1	ENAH MOUSE
16	253	6.0	802	1	MYH9 HUMAN
17	253	6.0	1938	1	BN1 YEAST
18	253	6.0	1953	1	MYH9 RAT
19	252.5	6.0	944	1	NUPI YEAST
20	252	6.0	1935	1	MYSS CYPCA
21	251	5.9	1997	1	SCPI RAT
22	251	5.9	1978	1	MYHE CHICK
23	251	5.9	3259	1	GOB1 HUMAN
24	249.5	5.9	1020	1	CF60 HUMAN
25	249.5	5.9	1960	1	MYH3 HUMAN
26	249	5.9	1961	1	MYH9 RAT
27	247.5	5.8	1938	1	MYHD HUMAN
28	246.5	5.8	1979	1	TRIA HUMAN
29	245	5.8	1972	1	MYHB HUMAN
30	245	5.8	2871	1	DESP HUMAN
31	244.5	5.8	1935	1	MYH7 PIG
32	244.5	5.8	2116	1	MYH2 DICDI
33	243.5	5.8	1935	1	MYH7 HUMAN

34	243	5.7	2611	1	BP1E MOUSE
35	242.5	5.7	1102	1	MYSC CHICK
36	242.5	5.7	1597	1	KTRO_MOUSE
37	241	5.7	1330	1	KYNI_VULVU
38	241	5.7	1940	1	MYH3 RAT
39	240.5	5.7	697	1	MFPI LYCES
40	240.5	5.7	1690	1	C190 DROME
41	240	5.7	1972	1	MYHB MOUSE
42	240	5.7	2054	1	ML8A HUMAN
43	239	5.6	1248	1	DIAL HUMAN
44	239	5.6	1938	1	MYH6_MOUSE
45	238.5	5.6	1727	1	ALM1 SCHPO

ALIGNMENTS

RESULT 1
ID MEAG6 HUMAN STANDARD: PRT: 804 AA.
AC O15320: O00169;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Meningioma-expressed antigen 6/11 (MEAG6) (MEAL1).
GN MGEA6 OR MGEA11.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A. (ISOFORMS MEAG6 AND MEAL1).
TISSUE=Meningioma;
RX MEDLINE=97472454; PubMed=9356211;
RA Heckel D., Brass N., Fischer U., Blin N., Steudel I., Tuereci O.,
RA Packler O., Zang K.D., Meese E.;
RT "cDNA cloning and chromosomal mapping of a predicted coiled-coil
proline-rich protein immunogenic in meningioma patients.";
RL Hum. Mol. Genet. 6:2031-2041(1997).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=MEAG6;
CC IsoId=O15320-1; Sequence=Displayed;
CC Name=MEAL1;
CC IsoId=O15320-2; Sequence=VSP 004322;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Widely expressed. Expressed in brain, muscle
and cranial skin.
CC -!- DISEASE: AUTOCANTIGEN FOUND IN BENIGN MENINGIOMA.
CC -!- CAUTION: At least 9 pseudogenes have been found on different
chromosomes, including chromosomes 2, 3, 6, 7, 9, 10, 12, 13 and
18.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U94780; AAB86593.1;
EMBL; U73682; AAB86589.1; ALT_INIT.
Genew; HGNC:7057; MGEA6.
MIM; 602132;
DR GO; GO:0008047; F:enzyme activator activity; TAS.
DR GO; GO:0005515; F:protein binding; TAS.
DR GO; GO:0005515; F:protein binding; TAS.
KW Antigen; Coiled coil; Alternative splicing
FT DOMAIN 87 273 COILED COIL (POTENTIAL).
FT DOMAIN 321 501 COILED COIL (POTENTIAL).
FT DOMAIN 553 558 POLY-GLY.
FT DOMAIN 502 804 PRO-RICH.
FT VARSPLIC 514 556 Missing (in isoform MEAL1).

```

FT CONFLICT 198 198 /FTID=VSP 004322.
FT CONFLICT 360 360 P -> Q (IN REF. 1; AAB85889).
FT CONFLICT 589 589 L -> P (IN REF. 1; AAB85889).
FT CONFLICT 594 594 F -> S (IN REF. 1; AAB85889).
FT CONFLICT 699 699 V -> I (IN REF. 1; AAB85889).
FT CONFLICT 738 738 G -> R (IN REF. 1; AAB85889).
SQ SEQUENCE 804 AA; 30943 MW; 3F088799E0883EDE CRC64;

Query Match 92.7%; Score 3923; DB 1; Length 804;
Best Local Similarity 93.1%; Pred. No. 3e-136;
Matches 750; Conservative 21; Mismatches 33; Indels 2; Gaps 2;

QY 1 MEEPGATPOYGLVLEELRRVVAALPESMRDENPYPPSELVCAAVIGFFVVLFLW 60
DB 1 MEEPGATPOYGLVLEELRRVVAALPEGMWPDNLYGFFWELVCAAVGFFVFLW 60

QY 61 RSFRSRSRLYVGRQKLGATSLGIEKCKLLEKPSLIQKEYEGYEVESSELSDFKE 120
DB 61 RSFRSRSRLYVGRQKLGATSLGIEKCKLLEKPSLIQKEYEGYEVESSELSDFKE 120

QY 121 AABEASLEATCEKLNRSSELEDETLCKDLKQEKSKHQSQODELMADISKISQLEDE 180
DB 121 AABEASLEATCEKLNRSSELEDETLCKDLKQEKSKHQSQODELMADISKISQLEDE 180

QY 121 AT-EASLSLEATCEKLNRSSELEDETLCKDLKQEKSKHQSQODELMADISKISQLEDE 179
DB 121 AT-EASLSLEATCEKLNRSSELEDETLCKDLKQEKSKHQSQODELMADISKISQLEDE 179

QY 181 SKSLKQIAIAKICTFKMSBERRAIAKDALNENSQLOTHKQIFQEAENVKGEVSE 240
DB 181 SKSLKQIAIAKICTFKMSBERRAIAKDALNENSQLOTHKQIFQEAENVKGEVSE 240

QY 180 SKSLKQIAIAKICTFKMSBERRAIAKDALNENSQLOTHKQIFQEAENVKGEVSE 238
DB 180 SKSLKQIAIAKICTFKMSBERRAIAKDALNENSQLOTHKQIFQEAENVKGEVSE 238

QY 241 LNKQITFEDSKVHAEQVNDKSHKTLTGLPMKQAAVLEEDTDDNLELVNSE 300
DB 241 LNKQITFEDSKVHAEQVNDKSHKTLTGLPMKQAAVLEEDTDDNLELVNSE 300

QY 239 LNKQITFEDSKVHAEQVNDKSHKTLTGLPMKQAAVLEEDTDDNLELVNSE 298
DB 239 LNKQITFEDSKVHAEQVNDKSHKTLTGLPMKQAAVLEEDTDDNLELVNSE 298

QY 301 SENGAYLDNPPGALKKLIHAALNASLKTLEGERNQIYQISEVDKTEELTEHKNQ 360
DB 301 SENGAYLDNPPGALKKLIHAALNASLKTLEGERNQIYQISEVDKTEELTEHKNQ 360

QY 299 SENGAYLDNPPGALKKLIHAALNASLKTLEGERNQIYQISEVDKTEELTEHKNQ 358
DB 299 SENGAYLDNPPGALKKLIHAALNASLKTLEGERNQIYQISEVDKTEELTEHKNQ 358

QY 361 TQASLOSENTHFENENQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 420
DB 361 TQASLOSENTHFENENQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 420

QY 359 TQASLOSENTHFENENQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 418
DB 359 TQASLOSENTHFENENQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 418

QY 421 EKISHATELETYKRAKDLLEELERTIHSYQQLIISHEKKAHDNWLAAARNLNDLR 480
DB 421 EKISHATELETYKRAKDLLEELERTIHSYQQLIISHEKKAHDNWLAAARNLNDLR 480

QY 419 EKISHATELETYKRAKDLLEELERTIHSYQQLIISHEKKAHDNWLAAARNLNDLR 478
DB 419 EKISHATELETYKRAKDLLEELERTIHSYQQLIISHEKKAHDNWLAAARNLNDLR 478

QY 481 KENAHNRQKLTETELKTELLEKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 540
DB 481 KENAHNRQKLTETELKTELLEKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 540

QY 479 KENAHNRQKLTETELKTELLEKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 538
DB 479 KENAHNRQKLTETELKTELLEKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 538

QY 541 TLLEGLRLSPLPFGGGRGSRGPNPLDHOITNERGESSCDRLTDPHAPSSTGSLSP 600
DB 541 TLLEGLRLSPLPFGGGRGSRGPNPLDHOITNERGESSCDRLTDPHAPSSTGSLSP 600

QY 539 TLLEGLRLSPLPFGGGRGSRGPNPLDHOITNERGESSCDRLTDPHAPSSTGSLSP 598
DB 539 TLLEGLRLSPLPFGGGRGSRGPNPLDHOITNERGESSCDRLTDPHAPSSTGSLSP 598

QY 601 WQDRRMFPPPGQSPDLSALPPQDRFCNSGRLSGPAELRSFNWPSLDKWDGMPSE 660
DB 601 WQDRRMFPPPGQSPDLSALPPQDRFCNSGRLSGPAELRSFNWPSLDKWDGMPSE 660

QY 599 WQDRRMFPPPGQSPDLSALPPQDRFCNSGRLSGPAELRSFNWPSLDKWDGMPSE 658
DB 599 WQDRRMFPPPGQSPDLSALPPQDRFCNSGRLSGPAELRSFNWPSLDKWDGMPSE 658

QY 661 MESSRNDTKDLGNLNVDSLSLPAENATGPGFVPPPLAPIRGFLFPVDARGPFLRRGPP 720
DB 661 MESSRNDTKDLGNLNVDSLSLPAENATGPGFVPPPLAPIRGFLFPVDARGPFLRRGPP 720

QY 659 MESSRNDTKDLGNLNVDSLSLPAENATGPGFVPPPLAPIRGFLFPVDARGPFLRRGPP 718
DB 659 MESSRNDTKDLGNLNVDSLSLPAENATGPGFVPPPLAPIRGFLFPVDARGPFLRRGPP 718

QY 721 FPPPPPGAMGASRDYFFPPFPFGPPPPAPFAMRNVPYPPRGGPPPLPPRPPFPFPPHSEG 780
DB 721 FPPPPPGAMGASRDYFFPPFPFGPPPPAPFAMRNVPYPPRGGPPPLPPRPPFPFPPHSEG 780

QY 719 FPPPPPGAMGASRDYFFPPFPFGPPPPAPFAMRNVPYPPRGGPPPLPPRPPFPFPPHSEG 778
DB 719 FPPPPPGAMGASRDYFFPPFPFGPPPPAPFAMRNVPYPPRGGPPPLPPRPPFPFPPHSEG 778

QY 781 RSEFSGGLIPSPNEPATEHPPEQQT 806
DB 779 RSEFSGGLIPSPNEPATEHPPEQQT 804

RESULT 2
MYH9_CHICK
ID MYH9_CHICK STANDARD; PRT; 1959 AA.

```

```

AC PI4105;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Myosin heavy chain, nonmuscle (Cellular myosin heavy chain) (NMNHC).
GN MYH9.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
OX [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Intestinal epithelium;
RC MEDLINE=9004668; PubMed=2813355;
RA Shohet R.V., Corti M.A., Kawamoto S., Preston Y.A., Brill D.A.,
RA Adelman R.S.;
RA "Cloning of the cDNA encoding the myosin heavy chain of a vertebrate
RT cellular myosin.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:7726-7730(1989).
CC -!- FUNCTION: Cellular myosin appears to play a role in cytokinesis,
CC cell shape, and specialized functions such as secretion and
CC capping.
CC -!- SUBUNIT: Myosin is a hexameric protein that consists of 2 heavy
CC chain subunits (MHC), 2 alkali light chain subunits (MLC) and 2
CC regulatory light chain subunits (MLC-2).
CC -!- DOMAIN: The rodlike tail sequence is highly repetitive, showing
CC cycles of a 28-residue repeat pattern composed of 4 heptapeptides,
CC characteristic for alpha-helical coiled coils.
CC -!- SIMILARITY: Contains 1 myosin-like globular head domain.
CC -!- SIMILARITY: Contains 1 IQ domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: M26510; AAA48974.1;
CC PIR: A33977; A33977.
CC HSP: P10587; IBR2.
CC InterPro: IPR000048; IQ region.
CC InterPro: IPR001609; Myosin_head.
CC InterPro: IPR004009; Myosin_N.
CC InterPro: IPR002928; Myosin_tail.
CC Pfam: PF00612; IQ; 1.
CC Pfam: PF00663; Myosin_head; 1.
CC Pfam: PF02736; Myosin_N; 1.
CC Pfam: PF01576; Myosin_tail; 1.
CC PRINTS: PR00193; MYOSINHEAVY.
CC ProDom: PD000355; myosin_head; 1.
CC SMART: SM00015; IQ; 1.
CC SMART: SM00242; MYSC; 1.
CC PROSITE: PS50096; IQ; 1.
CC Myosin; ATP-binding; Calmodulin-binding; Actin-binding;
KW Coiled coil; Multigene family. MYOSIN HEAD-LIKE.
FT DOMAIN 1 778
FT DOMAIN 779 808
FT DOMAIN 837 1925
FT NP_BIND 174 181
FT NP_BIND 654 676
FT DOMAIN 1959 AA; 226502 MW; A75C6086FD3A1A1 CRC64;
SQ SEQUENCE 1959 AA; 226502 MW; A75C6086FD3A1A1 CRC64;

Query Match 6.5%; Score 276.5; DB 1; Length 1959;
Best Local Similarity 24.5%; Pred. No. 0.0028;
Matches 133; Conservative 96; Mismatches 173; Indels 141; Gaps 26;

QY 59 LWRSEFVSRLVGRQKLGATSLGIEKCKLLEKPSLIQKEYEGYEVESSELSDFKE 118
DB 1299 LAKDFALSQLOQTQE-----LLQETRLKLSFSTKLQTE--DEKVALKELEE 1347

```


1650 1660 1670 1680 1690 1700 1710
GGAGACCTATAGAAACCGAGCCAAAGATCTTGAAGAGGATGGAGAGACTATTTCATCTTATCAAGGGCA
-----ACTGACCACTATTTGCTCAATATAATTAATAGAGAGTATTCCTTTGGGTTTTTTTTTGAA
660 670 680 690 700 710
1720 1730 1740 1750 1760 1770 1780
GATTATTCCCATGCGAAGAAAGACATCATATGCTTGGCAGCTCGGAATGCTGAAAGAAACCTCAATGA
AAATGT-----TGAAGAGGGAGTGTGTTTAAATGAAATTTTTTCCAAACGAAAGACCACAAATTTCTTAG
720 730 740 750 760 770 780
1790 1800 1810 1820 1830 1840 1850
TTTAGGAAAGAAATGCTCACACACAGACAAATAATTAATCT--GAAACAGAGCTTAAATTTGAATTTTAGA
TATCGGTCCGCGAATCATTTGTACCGTCAAAAAAAATTTGGTTTAAAAACATGATCTCTGAAGTTAAAAA
790 800 810 820 830 840 850
1860 1870 1880 1890 1900 1910 1920
AAAAGATCCTTATGCACTCGATGTTCCAAATACAGCATTTGGCAGAGAGCATTTCCCATATGTCCTCCAC
ATCCGCTGAAATTTGGTCGATTTCTCGGTTATTCA---AGCAAAATGATCCGCAAGCACTGATGTTGGCG
860 870 880 890 900 910 920
1930 1940 1950 1960 1970 1980 1990
ATTGGGTTGCTTCTATCTGAAACAGAGCTTTCTCTCTCTCCAACTTTGTTGAGGGTCCACTCAGACT
AAAAGAACGAAACACGCGAAGATAGCCCTCAATGTCCAACGTGGAGATCGAGCGCTTGGTTCCGCAAGACA
930 940 950 960 970 980 990
2000 2010 2020 2030 2040 2050 2060
CTCACCTTTCTCCAGGGGGAGAGAGAGCTCACAGGCCCGAGGAATCCTCTGGACCATCAGATTAC
AGAGGAGCCCTTTTAAAGCCCGCGACGGAACCTGAAGAGGAGCTTCAGATAGCTGAAAGCGT
1000 1010 1020 1030 1040 1050 X
2070 2080 2090 2100
CAATGAAGAGGAGATCAAGCTGTGATAGTTAACCGA

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2004, 07:50:53 ; Search time 84 Seconds
(without alignments)
3027.472 Million cell updates/sec

Title: US-09-744-314-6
Perfect score: 4231
Sequence: 1 MEEFGATPQYLGLVLELR.....GLIPSNPEATBEPFQQT 806

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- SPREMBL 25: *
- 1: sp_archaea: *
- 2: sp_bacteria: *
- 3: sp_fungi: *
- 4: sp_human: *
- 5: sp_invertebrate: *
- 6: sp_mammal: *
- 7: sp_mhc: *
- 8: sp_organelle: *
- 9: sp_phage: *
- 10: sp_plant: *
- 11: sp_rodent: *
- 12: sp_virus: *
- 13: sp_vertebrate: *
- 14: sp_unclassified: *
- 15: sp_rvirus: *
- 16: sp_bacterioph: *
- 17: sp_archaeap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3812	90.1	775	4 Q8IX93	Q8IX93 homo sapien
2	3652.5	86.3	788	4 Q86T56	Q86T56 homo sapien
3	3643	86.1	771	4 Q8IX92	Q8IX92 homo sapien
4	3626	85.7	777	4 Q95046	Q95046 homo sapien
5	3529	83.4	811	4 Q8NE23	Q8NE23 homo sapien
6	3312	78.3	732	4 Q86U72	Q86U72 homo sapien
7	3253	76.9	754	4 Q86R76	Q86R76 homo sapien
8	3235	76.5	745	4 Q8NE38	Q8NE38 homo sapien
9	2921.5	69.0	779	11 Q8R311	Q8R311 mus musculus
10	2317.5	54.8	582	11 Q8CIE3	Q8CIE3 mus musculus
11	1605	37.9	370	4 Q8IX94	Q8IX94 homo sapien
12	1542.5	36.5	365	4 Q8N90	Q8N90 homo sapien
13	1242.5	29.4	691	13 Q7ZU98	Q7ZU98 brachydanio
14	766	18.1	1193	4 Q82580	Q82580 homo sapien
15	605.5	14.3	407	13 Q7S274	Q7S274 xenopus lae
16	546	12.9	158	4 Q8IX95	Q8IX95 homo sapien

17	502	11.9	461	4 Q86Y60	Q86Y60 homo sapien
18	345	8.2	1430	5 Q9VMA7	Q9VMA7 drosophila
19	305	7.2	1326	4 Q9P219	Q9P219 homo sapien
20	286	6.8	1239	5 Q07569	Q07569 entameba h
21	285.5	6.7	1257	5 Q9VMA8	Q9VMA8 drosophila
22	276	6.5	1999	11 Q63731	Q63731 rattus norv
23	274.5	6.5	1331	4 Q86T83	Q86T83 homo sapien
24	271.5	6.4	1784	13 Q90Y85	Q90Y85 gallus gall
25	271.5	6.4	1743	5 Q96063	Q96063 dugesia jap
26	266	6.3	1427	11 Q80XQ1	Q80XQ1 mus musculu
27	266	6.3	1790	3 Q07380	Q07380 saccharomyc
28	264.5	6.3	1762	10 Q94DC2	Q94DC2 oryza sativ
29	263.5	6.2	1833	11 Q80T68	Q80T68 mus musculu
30	263	6.2	1323	4 Q8N625	Q8N625 homo sapien
31	263	6.2	1421	4 Q9P257	Q9P257 homo sapien
32	263	6.2	1636	4 Q9H357	Q9H357 homo sapien
33	262	6.2	1958	5 Q96062	Q96062 dugesia jap
34	261.5	6.2	742	5 Q813B2	Q813B2 plasmodium
35	261	6.2	1940	5 Q9U7E3	Q9U7E3 pecten maxi
36	260.5	6.2	2612	5 Q815X5	Q815X5 plasmodium
37	258.5	6.1	895	11 Q811I2	Q811I2 mus musculu
38	258.5	6.1	2007	13 Q02015	Q02015 gallus gall
39	258	6.1	2011	5 Q8MNC2	Q8MNC2 drosophila
40	258	6.1	2017	5 Q94992	Q94992 drosophila
41	258	6.1	2056	5 Q9A0W8	Q9A0W8 drosophila
42	258	6.1	2057	5 Q94987	Q94987 drosophila
43	257.5	6.1	1837	3 Q74424	Q74424 schizosacch
44	257	6.1	1831	10 Q7XN11	Q7XN11 oryza sativ
45	257	6.1	1953	3 Q13450	Q13450 saccharomyc

ALIGNMENTS

RESULT 1

Q8IX93 PRELIMINARY; PRT; 775 AA.
AC Q8IX93
DT 01-MAR-2003 (EMBLrel. 23, Created)
DT 01-MAR-2003 (EMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (EMBLrel. 23, Last annotation update)
DE CTAGE-5A protein.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA User: D., Schaefer D., Eichmuller S.;
RT "CTAGE-family members"
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF338233; A017610.1; --
SQ SEQUENCE 775 AA; 87970 MW; DF5CB0BBCEA390C2 CRC64;

Query Match	90.1%;	Score 3812;	DB 4;	Length 775;
Best Local Similarity	93.7%;	Pred. No. 1.5e-166;		
Matches 728;	Conservative 18;	Mismatches 29;	Indels 2;	Gaps 2;
QY	30	MRPDNPYGFPPSELVCAAVGFFVLLPLNRSFVSRLVYVGEQKLGATLSGLIEEK	89	
Db	1	MRPDNLXGFPWELVCAAVGFFVLLPLNRSFVSRLVYVGEQKLGATLSGLIEEK	60	
QY	90	CKLLEKFSLIQYEGYEVESSELDASPEKZAAEARSLEATCEKLNRSNSELEDEILCL	149	
Db	61	SKLLEKFSLVQYEGYEVESSELDASPEKZAAEARSLEATCEKLNRSNSELEDEILCL	119	
QY	150	EKDLKOEKSKHQDELMDADISKSIOSEDESKSKQAIAEKIKCTPKMSEERRATAI	209	
Db	120	EKELKEEKSQDELMDADISKRICQSEDESKSKQVAEAKMTFKIQQMNEELKTAI	179	
QY	210	KDALNENSOLOTSKOLFOQBAEVNKGVESELNKKITFEDSKVHAEQVLDNKHITKL	269	
Db	180	KDALNENSOLOTSKOLFOQBAEVNKGVESELNKKITFEDSKVHAEQVLDNKHITKL	238	

QY 270 TGHLPMMKQAAVLBEDTTDDNLELVNSENSENGAYLDNPPGKALKKLIHAAKLNASLK 329
 Db 239 TERLLKKOKWAAMIGDITDDNLELVNSENSENGAYLDNPPGKALKKLIHAAKLNASLK 298
 QY 330 TLEGERNQIYIOLSEVDKTKBELTEHIKNLOTCQASLOSENTHFENENKQKQKQKQKQK 389
 Db 299 TLEGERNQIYIOLSEVDKTKBELTEHIKNLOTCQASLOSENTHFENENKQKQKQKQKQK 358
 QY 390 LYQENEMKLRKLTVEENYRLEKEKLSKVDEKISHATEELETYKRAKDLSEELERTIH 449
 Db 359 LYQENEMKLRKLTVEENYRLEKEKLSKVDEKISHATEELETYKRAKDLSEELERTIH 418
 QY 450 SYQGQIISHEKKAHDNLAARNAERNDLRKENAHNRKLTETELKPELLEKDPYALDV 509
 Db 419 SYQGQIISHEKKAHDNLAARNAERNDLRKENAHNRKLTETELKPELLEKDPYALDV 478
 QY 510 PNTAFORSHSPGSPPLGWSSETRAFLSPPTLLGEPRLSLPLPGGGGGRSGRSGNPLD 569
 Db 479 PNTAFORSHSPGSPPLGWSSETRAFLSPPTLLGEPRLSLPLPGGGGGRSGRSGNPLD 538
 QY 570 HOITNERGESSCDRLTDHRAPSDTSGLSPWDQRRMMFPFGGSGYSDSALPPQQRDF 629
 Db 539 HOITNERGESSCDRLTDHRAPSDTSGLSPWDQRRMMFPFGGSGYSDSALPPQQRDF 598
 QY 630 CNSGSLSPGPAELRSNMPSLDKMDGSMPESSSRNDTKDGLNLPDSSLPNAENAT 689
 Db 599 CNSGSLSPGPAELRSNMPSLDKMDGSMPESSSRNDTKDGLNLPDSSLPNAENAT 658
 QY 690 GGFVPPPLAPIRGPIFVDAGPFLRGPFPPLPPPPGAMFGASRDYPPRDRFPGPFPAP 749
 Db 659 GGFVPPPLAPIRGPIFVDAGPFLRGPFPPLPPPPGAMFGASRDYPPRDRFPGPFPAP 718
 QY 750 FANRNVYPRGPPYLPFRPGFFPPPHSGRSEPPSGLIIPPSNEPATEHPPEQOET 806
 Db 719 FANRNVYPRGPPYLPFRPGFFPPPHSGRSEPPSGLIIPPSNEPATEHPPEQOET 775

RESULT: 2
 Q86TF6 PRELIMINARY; PRT; 788 AA.
 AC Q86TF6;
 DT 01-JUN-2003 (T-EMBLrel. 24, Created)
 DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
 DE Similar to CTAGE-2 (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=B-cell;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeng B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Donald M.P., Casavant T.L., Schetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalek U., Smalley D.E., Schnerk A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=B-cell;
 RA Strausberg R.;
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC051363; AAHS1363.1;
 FT NON TER 1
 SQ SEQUENCE 788 AA; 89147 MW; 243CA46DFD346686 CRC64;
 Query Match 86.3%; Score 3652.5; DB 4; Length 788;
 Best Local Similarity 93.1%; Pred. No. 3e-159;
 Matches 701; Conservative 18; Mismatches 27; Indels 7; Gaps 3;
 QY 59 LMRFRVRRLVYVREKQKLGATLSGLIBKCKLLEKFKLSLIQKEYEGYEVESLEDSPE 118
 Db 38 LAQHCLKYRSLVYVREKKLALMLSLGLIBKCKLLEKFKLSLIQKEYEGYEVESLEDSPE 97
 QY 119 KBAABEARSLE-ATCEKLNRSNSELDEILCLEKDKQKSKSHSQODELMADISKS 173
 Db 98 KEAT-EAQSLEVENQWATCEKLNRSNSELDEILCLEKDKQKSKSHSQODELMADISKR 156
 QY 174 IQSLEDESKLSQIAEAKTICTFKMSBERRAIAIKDALNENSQLOTHKQLPQOEAEV 233
 Db 157 IQSLEDESKLSQVAAEKMTFKI FQWNEERLKAIAIKDALNENSQLOTHKQLPQOEAEV 215
 QY 234 WKGEVSELNKKITFDSKVHAEQVLDNKHENIKITLTHLPMKQAAVLBEDTTDDN 293
 Db 216 WKEQVSELNKKITFDSKVHAEQVLDNKHENIKITLTHLPMKQAAVLBEDTTDDN 275
 QY 294 ELEVNSENGAYLDNPPGKALKKLIHAAKLNASLKTEGERNQIYIOLSEVDKTKBELT 353
 Db 276 ELEVNSENGAYLDNPPGKALKKLIHAAKLNASLKTEGERNQIYIOLSEVDKTKBELT 335
 QY 354 EHKNLQTOASQSENTHFENENKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQK 413
 Db 336 EHKNLQTOASQSENTHFENENKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQK 395
 QY 414 EKLKVDKISHATEELETYKRAKDLSEELERTIHSHKKAHDNLAARNAE 473
 Db 396 EKLKVDKISHATEELETYKRAKDLSEELERTIHSHKKAHDNLAARNAE 455
 QY 474 RNLDLREKNAHNRKLTETELKPELLEKDPYALDVNTAFGRHSPYGPSPGWSSET 533
 Db 456 RNLDLREKNAHNRKLTETELKPELLEKDPYALDVNTAFGRHSPYGPSPGWSSET 515
 QY 534 RAFLSPPTLLGEPRLSLPLPGGGGGRSGRSGNPLDTPHRAPSD 593
 Db 516 RAFLSPPTLLGEPRLSLPLPGGGGGRSGRSGNPLDTPHRAPSD 575
 QY 594 TGSLSPPWDQRRMMFPFGGSGYSDSALPPQQRDFCSNSGRSLSPGPAELRSNMPSLDKM 653
 Db 576 TGSLSPPWDQRRMMFPFGGSGYSDSALPPQQRDFCSNSGRSLSPGPAELRSNMPSLDKM 635
 QY 654 DGSMPSESSSRNDTKDGLNLPDSSLPNAENATGPGVPPPLAPIRGPIFVDAGP 713
 Db 636 DGSMPSESSSRNDTKDGLNLPDSSLPNAENATGPGVPPPLAPIRGPIFVDAGP 695
 QY 714 FLRRGPPFPFPFGAMFGASRDYPPRDRFPGPFPAPFAMNVYPRGFFPYLPFRPGFFP 773
 Db 696 FLRRGPPFPFPFGAMFGASRDYPPRDRFPGPFPAPFAMNVYPRGFFPYLPFRPGFFP 755
 QY 774 PPPHSGRSEPPSGLIIPPSNEPATEHPPEQOET 806
 Db 756 PPPHSGRSEPPSGLIIPPSNEPATEHPPEQOET 788

RESULT 3
 Q8IX92 PRELIMINARY; PRT; 771 AA.
 AC Q8IX92;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2004, 07:50:53 ; Search time 21 Seconds
(without alignments)
3691.922 Million cell updates/sec

Title: US-09-744-314-6
Perfect score: 4231
Sequence: 1 MEEPGATPQYGLGLVLELR.....GLIPPSNEPATEHPPEPQQT 806

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	286	6.8	2139	2 T18296	myosin heavy chain
2	276.5	6.5	1959	1 A33977	myosin heavy chain
3	276	6.5	1999	1 S21801	myosin heavy chain
4	274.5	6.5	1130	2 T34081	hypothetical prote
5	271.5	6.4	1875	2 S38173	myosin-like protei
6	270	6.4	1300	2 I53799	CGI protein - huma
7	270	6.4	1356	2 S32763	kinectin 1 - huma
8	270	6.4	1410	1 A57013	early endosome ant
9	268	6.3	1957	2 T38077	hypothetical coile
10	266.5	6.3	1976	2 A59252	myosin heavy chain
11	266	6.3	1790	2 S67593	transport protein
12	258.5	6.1	2007	1 B43402	myosin heavy chain
13	258	6.1	2017	1 A36014	myosin heavy chain
14	258	6.1	2057	2 S61477	myosin II heavy ch
15	257.5	6.1	1837	2 T41023	probable nuclear p
16	257	6.1	1961	1 A61231	myosin heavy chain
17	256.5	6.1	1931	2 A86234	slow myosin heavy
18	256	6.1	841	2 A86188	hypothetical prote
19	254.5	6.0	1939	2 T18372	repeat organellar
20	254	6.0	2253	2 T30336	nuclear/mitotic ap
21	253	6.0	1938	1 A40997	myosin heavy chain
22	253	6.0	1953	2 S63244	BN11 protein - yea
23	252.5	6.0	944	2 S26710	spindle pole body
24	252	6.0	1940	2 A59287	myosin heavy chain
25	252	6.0	2020	2 T21174	hypothetical prote
26	251	5.9	946	2 S28061	SCP1 protein - rat
27	251	5.9	1979	1 S03166	myosin heavy chain
28	251	5.9	2288	2 T29999	hypothetical prote
29	251	5.9	3225	2 I52300	giantin - human

30 251 5.9 3259 1 A56539 giantin - human
31 250 5.9 1738 2 T14867 interaptin - slime
32 249.5 5.9 2954 2 T14156 kinesin-related pr
33 249 5.9 1992 2 A47237 myosin heavy chain
34 248 5.9 1538 2 T29095 cardiac muscle fac
35 247.5 5.8 746 2 T47237 myosin II heavy ch
36 247 5.8 1313 2 F96673 hypothetical prote
37 246.5 5.8 1164 2 T24806 hypothetical prote
38 246 5.8 1388 2 S74245 serine/threonine-s
39 246 5.8 1909 2 A45592 liver stage antige
40 245 5.8 2677 2 A38194 desmoplakin I - hu
41 244.5 5.8 1175 2 D35815 myosin heavy chain
42 244.5 5.8 1201 2 B35815 myosin heavy chain
43 244.5 5.8 1935 2 A59286 myosin heavy chain
44 244.5 5.8 1964 2 A59282 nonmuscle myosin I
45 244.5 5.8 1974 2 T16703 hypothetical prote

ALIGNMENTS

RESULT 1

T18296

myosin heavy chain - Entamoeba histolytica

C;Species: Entamoeba histolytica

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 08-Sep-2000

C;Accession: T18296

R;Guillen, N.

submitted to the EMBL Data Library, February 1997

A;Reference number: Z18865

A;Accession: T18296

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-2139 <GUI>

A;Cross-references: EMBL:L03534; NID:G1850912; PID:G1850913; PIDN:AAB48065.1

C;Genetics:

A;Gene: mhca

C;Superfamily: myosin heavy chain; myosin motor domain homology

F;91-780/Domain: myosin motor domain homology <MVO>

Query Match 6.8%; Score 286; DB 2; Length 2139;

Best Local Similarity 21.2%; Pred. No. 0.0068;

Matches 127; Conservative 105; Mismatches 173; Indels 194; Gaps 21;

QY	81	TLSGLTEEEKKLEKFP--SLIQKEVEGEVSESSLESDASFEKEAABEARS-LEAT----	CE 133
DB	977	TLNATVKKDKITIAEQESIDEKEDITKLGDKILKEEKDDLEQDRAVDVATKDDIAK 1036	
QY	134	KLNRNSELE---DEILCEKDLKQKSHSQODELMADISKISIQSLEDESKLSQIAE 190	
DB	1037	KLNKITIECEDAKDEIAKLEQLEDEENKKN---DITNELQQTQLKLGTEKSLAAQVAA 1093	
QY	191	AKIICKTFKMSERRAIAIKDALN---ENSQIQTSH-----	223
DB	1094	TK-----KASDER-----DTLSQNLNEKLTTKTKTKADLEKKISGLKQDYEDLED 1141	
QY	224	-----KQLFQQAEEVNVKGEVSELNKQKITTFDESKVH 254	
DB	1142	DKNKIEGLRNAQKIKELDDITTKGADVSYQLQKQKEEVESQIAKMQEKEAIGNDKVN 1201	
QY	255	AEQVLNDKNEHKITLGHLPMMKDDQAAVLEEDT-----TDDDNLELVNS 299	
DB	1202	KEKTIKEKELEIISLOEKL-----DETEVEKEDAEKKKKEKEMKALQEBKENVESKNS 1257	
QY	300	ESENGAYLDNPPPGALKKL-----IHAAKLNASLTKTLEGERNQIY-----IQ 341	
DB	1258	TERDKKLDNLDKTDQKKDDMTADNEKAKAKADLEAQLNEVDNHEKAVADAELLNKK 1317	
QY	342	LSEVDKTEBELTHIKNLQTOQASLQSENTHFNENQKLOQKL-----KV 386	
DB	1318	KAQSDKELNSLKAELEALTAKKSVESKNSKDSSEKAAUGSEIIDQANEKLNQADLRKA 1377	
QY	387	MTELYQENEMKLN-----HRKLTIV-----EEN-YRLEKEEK-LSKVDEKI 423	

```
Db 1378 TADLOEANEKAEVAQDRKLVADNKKMTKTLEEIKARDEENTYKVENYKVLKREKADL 1437
Qy 424 SHATELETYR-----KRAKDLEBEELERTHYSQOQIISHEK----- 460
Db 1438 BRANENWLDKDRMNKQKQVGLSGELKETKDKLNAIAEKDSIFTAKKQSDADLEELN 1497
Qy 461 -----KAHDNWLAAARN-----AERNLNDLRKENAHNRQKUTETELKPELLEKDP 504
Db 1498 KTVSEHDEVAKLNTQITKLTDRNQSABEELNELRSKADKOKKKTISEEQVNELESRP 1556

RESULT 2
A33977 myosin heavy chain, nonmuscle - chicken
N:Contains: myosin ATPase (EC 3.6.4.1)
C:Species: Gallus gallus (chicken)
C>Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text_change 19-Apr-2002
C:Accession: A33977; S06116; A43422
R:Shohet, R.V.; Conti, M.A.; Kawamoto, S.; Preston, Y.A.; Brill, D.A.; Adelstein, R.S.
Proc. Natl. Acad. Sci. U.S.A. 86, 7726-7730, 1989
A>Title: Cloning of the cDNA encoding the myosin heavy chain of a vertebrate cellular my
A:Reference number: A33977; MUID:90046668; PMID:2813355
A:Accession: A33977
A:Molecule type: mRNA
A:Residues: 1-1959 <SHO>
A:Cross-references: GB:M26510; NID:G212382; PIDN:AAA48974.1; PID:G212383
R:Katsuragawa, Y.; Yanagisawa, M.; Inoue, A.; Maeaki, T.
Eur. J. Biochem. 184, 611-616, 1989
A>Title: Two distinct nonmuscle myosin-heavy-chain mRNAs are differentially expressed in
s.
A:Reference number: S06116; MUID:90032648; PMID:2806244
A:Accession: S06116
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 716-1008 <KAT>
A:Cross-references: GB:X17589
A>Note: this translation is not annotated in GenBank entry GGMCFMHA, release 114
R:Hodge, T.P.; Cross, R.; Kendrick-Jones, J.
J. Cell Biol. 118, 1085-1095, 1992
A>Title: Role of the COOH-terminal nonhelical tailpiece in the assembly of a vertebrate
A:Reference number: A43422; MUID:92381096; PMID:1512291
A:Accession: A43422
A>Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1900-1959 <HOD>
A:Experimental source: brush border
A>Note: sequence extracted from NCBI backbone (NCBI:P111947)
A:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; nucleotid
F:84-764/Domain: myosin motor domain homology <MMOT>
F:174-181/Region: nucleotide-binding motif A (P-loop)
F:552-565/Region: actin binding #status predicted
F:626-640/Region: actin binding #status predicted
F:837-1936/Domain: coiled coil #status predicted <COI>
F:837-1277/Region: S2
F:1278-1959/Region: light meromyosin
F:1937-1959/Domain: carboxyl-terminal <CET>
F:125/Modified site: N6,N6-trimethyllysine (lys) #status predicted
F:180/Binding site: ATP (lys) #status predicted
F:694,704/Active site: Cys #status predicted

Query Match 6.5%; Score 276.5; DB 1; Length 1959;
Best Local Similarity 24.5%; Pred. No. 0.0014;
Matches 133; Conservative 96; Mismatches 173; Indels 141; Gaps 26;

Qy 59 LWRGFRVRSRLVYGRQKLGATSLGIEKCKLEKFKSLIQEYEGYVESLSLEDAFPE 118
Db 1299 LAKDFSALESQLOQTQ-----LLQETRLKLSFKSLKQTE--DEKNALKEQLEE 1347
Qy 119 KEAEERARSLEATCKLNR-----SNSLEDEILCLE-----KDL----- 153
Db 1348 EERAK--RNLEKQISVLQQQAVARKKMDGLGCLTAEBANKKLQKLESILQRIEYKI 1405
```

```
Qy 154 -----KOEKSK---HSQODELMADIS---KSIOLEDESKLSQTAIAKILICTFPMSEE 203
Db 1406 AAYDKLEKTRILQOELDDIAVDLDHQRTVSNLEKKQKKFPQLLAEEKNI--SAKYAE 1463
Qy 204 R-----RAIAIKDALNENSQLOTSKQLFQOEAEVWKGVESELNKKKIT 247
Db 1464 RDRAEAEAREKETKALSALARAEAEAEQKAELERVNKO--FRTEMEDLMSSKDDVGKSVHE 1522
Qy 248 FEDSKVHAGVLDNKENHIKTLTGHLPMKQDAVLEED--TTDDNLELVNSBSENGA 305
Db 1523 LEKAGALEQQOQVE-----MKTQLEEELELOATEDAKLELVNQQAAMKAQ 1568
Qy 306 Y-----LDNPPKGALKLIHAAK-----LNASLKTLEGERNOIYQLSE 344
Db 1569 FDRDLGRDEQNEKEKQLIRQVREMEVELEDERKORSIAVAARKKLELDLKLDSHIDI 1628
Qy 345 VDKTKSELTTHIKNLQTO-----QASIQSNTHEP-----NENKLOQKLKWTETLYQEN 394
Db 1629 ANKNRDEAIKHVRKLOAQKMDYMRLEEDTSTREILAQAENEK---KLKSW-----EA 1680
Qy 395 EM-KLHKLKTVENRYLEKEEKLKSKYDEKISH-----ATELETYRKRAKLESELER 446
Db 1681 EMIQLOEELAAERAKRQAOQERDELADETFANSSGKALAMEEKRLAELARIAQLESELEE 1740
Qy 447 TIHSYQG--QIISHE-KKAH---DNMLAARNLNDLRKENAHNRQKUTETELKPELL 500
Db 1741 E-----QGNTEIINDRLKKANLQIDQNMADLNABRS-NACKNENARQOMERQNKELKLQ 1795
Qy 501 EKD 503
Db 1796 EME 1798

RESULT 3
S21801 myosin heavy chain, neuronal [similarity] - rat
N:Alternate names: myosin II
N:Contains: myosin ATPase (EC 3.6.4.1)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text_change 19-Apr-2002
C:Accession: S21801; PNC0013; S18134
R:Sun, W.; Chantler, P.D.
J. Mol. Biol. 224, 1185-1193, 1992
A>Title: Cloning of the cDNA encoding a neuronal myosin heavy chain from mammalian brain
A:Reference number: S21801; MUID:92235856; PMID:1569576
A:Accession: S21801
A:Molecule type: mRNA
A:Residues: 1-1999 <SUN>
A:Cross-references: EMBL:X62659
R:Sun, W.; Chantler, P.D.
Biochem. Biophys. Res. Commun. 175, 244-249, 1991
A>Title: A unique cellular myosin II exhibiting differential expression in the cerebral c
A:Reference number: PNC0013; MUID:91151356; PMID:1998509
A:Accession: PNC0013
A:Molecule type: mRNA
A:Residues: 1914-1998, 'I' <SU2>
A:Experimental source: brain
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; nucleotide
F:84-763/Domain: myosin motor domain homology <MMOT>
F:174-181/Region: nucleotide-binding motif A (P-loop)
F:541-575/Region: actin binding #status predicted
F:653-675/Region: actin binding #status predicted
F:836-1959/Domain: coiled coil #status predicted <COI>
F:836-1276/Region: S2
F:1277-1999/Region: light meromyosin
F:125/Modified site: N6,N6-trimethyllysine (lys) #status predicted
F:180/Binding site: ATP (lys) #status predicted
F:693,703/Active site: Cys #status predicted
F:1916/Binding site: phosphate (Ser) (covalent) #status predicted
F:1943/Binding site: phosphate (Ser) (covalent) #status predicted
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2004, 07:50:53 ; Search time 60 Seconds
(without alignments)
3795.555 Million cell updates/sec

Title: US-09-744-314-6
Perfect score: 4231
Sequence: 1 MEEPGATPQYLGVLVLELR.....GLIPSPNEPATEHPEPQOET 806

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues
Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A Geneseq 29Jan04:*
- 1: Geneseqp1980s:*
 - 2: Geneseqp1990s:*
 - 3: Geneseqp2000s:*
 - 4: Geneseqp2001s:*
 - 5: Geneseqp2002s:*
 - 6: Geneseqp2003as:*
 - 7: Geneseqp2003bs:*
 - 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4231	100.0	806	3	AAY77574 Human cyt
2	3946.5	93.3	763	6	ABG74698 Human CGD
3	3897	92.1	804	6	ABG74685 Human CGD
4	3782	89.4	806	6	ABO07125 Novel hum
5	3740	88.4	881	4	ABG05280 Novel hum
6	3712	87.7	881	4	ABG20258 Novel hum
7	3705	87.6	760	6	ABO14717 Novel hum
8	3689.5	87.2	776	5	ABG4984 Human PRO
9	3689.5	87.2	776	5	ABG5590 Human ang
10	3689.5	87.2	776	5	ADD10625 Human sec
11	3689.5	87.2	776	5	ADD11585 Human sec
12	3689.5	87.2	776	5	ADD37378 Human sec
13	3689.5	87.2	776	8	ADE41586 Human sec
14	3674.5	86.8	776	4	ABG05279 Novel hum
15	3635	85.9	777	4	ABG20257 Novel hum
16	3634	85.9	777	4	AAH18372 Peptide #
17	3634	85.9	777	4	ABB37406 Peptide #
18	3634	85.9	777	4	AAH18372 Peptide #
19	3634	85.9	777	4	ABB32158 Peptide #
20	3634	85.9	777	4	ABG52211 Human liv
21	3634	85.9	777	5	ABG40185 Human pep
22	3626	85.7	777	4	AAH18368 Peptide #
23	3626	85.7	777	4	ABB37401 Peptide #
24	3626	85.7	777	4	AAH18368 Peptide #
25	3626	85.7	777	4	ABB32152 Peptide #

26	3626	85.7	777	4	ABG22697 Protein #
27	3626	85.7	777	4	AAH58083 Human bra
28	3626	85.7	777	4	ABG52207 Human liv
29	3626	85.7	777	4	AAH05968 Peptide #
30	3626	85.7	777	5	ABG40178 Human pep
31	3595	85.0	804	6	ABO07126 Novel hum
32	3593.5	84.9	1369	2	AAH24788 Human sec
33	3593.5	84.9	1369	4	AAH39009 Human sec
34	3593.5	84.9	1369	5	ABG55718 Human pol
35	3574.5	84.5	1086	4	ABG20256 Novel hum
36	3253	76.9	754	4	AAH70884 Human CTA
37	3190	75.4	798	6	ABJ19344 NOXV rela
38	3110	73.5	778	6	ABJ19345 NOXV rela
39	767	18.1	785	7	ADC37283 Nuclear f
40	766	18.1	1124	6	ABU70418 Human adi
41	766	18.1	1124	7	ADC37285 Nuclear f
42	766	18.1	1193	4	AAH25602 Human pro
43	766	18.1	1193	5	ABG61824 Prostate
44	766	18.1	1907	5	ABB82127 Human TAN
45	766	18.1	1907	7	ADE40158 Human NOV

ALIGNMENTS

RESULT 1
AAY77574
ID AAY77574 standard; protein; 806 AA.

XX AC AAY77574;
XX AC AAY77574;
XX 08-MAY-2000 (first entry)
XX Human cytoskeletal protein (HCVT) (clone 3768043).
XX Human cytoskeletal protein; HCVT; cell proliferation; immunological;
KW reproductive; nervous disorder; cytostatic; immunosuppressive; anti-HIV;
KW anti-diabetic; antiarteriosclerotic; dermatological; anti-inflammatory;
KW anti-infertility; vasotropic; cardiant.
XX Homo sapiens.
XX WO200006730-A2.
XX 10-FEB-2000.
XX 30-JUL-1999; 99WO-US017167.
XX 31-JUL-1998; 98US-0155185P.
XX 04-AUG-1998; 98US-0150081P.
XX 19-AUG-1998; 98US-0155228P.
XX (INCY-) INCYTE PHARM INC.
XX Bandman O, Tang YT, Yue H, Corley NC, Guegler KJ, Azimzai Y;
XX Patterson C, Lal P, Baughn MR;
XX WPI; 2000-195297/17.
XX N-PSDB; AAZ58979.
XX Human cytoskeletal proteins useful for diagnosing, treating preventing
XX cell proliferative, immunological, reproductive, developmental and
XX nervous disorders.
XX Claim 1; Page 69-70; 84pp; English.
XX The invention provides human cytoskeletal proteins (HCVT) and nucleic
XX acids encoding the proteins. The HCVT polypeptides can be expressed by
XX standard recombinant methodology. The HCVT polypeptides, modulators and
XX antibodies are useful for treating or preventing a disorder associated
XX with decreased and increased expression or activity of HCVT in mammals.
XX The polypeptides are also useful for diagnosing HCVT activity disorders
XX such as cell proliferative, immunological, reproductive, developmental

CC and nervous disorders. Sequences AAY77569-576 represent HCYT polypeptides

XX Sequence 806 AA;

Query Match 100.0%; Score 4231; DB 3; Length 806;
Best Local Similarity 100.0%; Pred. No. 8.1e-237;
Matches 806; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEEPGATPOPYLGVLVEELRRVVAALPESMRDENPYGFPSELVVCAAVIGFFVLLFLW 60
DB 1 MEEPGATPOPYLGVLVEELRRVVAALPESMRDENPYGFPSELVVCAAVIGFFVLLFLW 60

QY 61 RSFRVSRLYVGRQKLGATLSGLIEEKCKLLEKFSLIQKEYEGYEVESSELESPEKE 120
DB 61 RSFRVSRLYVGRQKLGATLSGLIEEKCKLLEKFSLIQKEYEGYEVESSELESPEKE 120

QY 121 AAEARSLEATCEKLNRSNSELDEILCLEKDKQKSKHSDQDELMADISKSQISLEDE 180
DB 121 AAEARSLEATCEKLNRSNSELDEILCLEKDKQKSKHSDQDELMADISKSQISLEDE 180

QY 181 SKSLASQIAEAKIICTFKMSERRAIAIKDALNENSQLOTSHKOLFQEAENVKGEVSE 240
DB 181 SKSLASQIAEAKIICTFKMSERRAIAIKDALNENSQLOTSHKOLFQEAENVKGEVSE 240

QY 241 LNKQKITPDSKVHAEQVLDNKENHILKTLGHLPMKQQAIVLEDDTTDDNLELVNSE 300
DB 241 LNKQKITPDSKVHAEQVLDNKENHILKTLGHLPMKQQAIVLEDDTTDDNLELVNSE 300

QY 301 SENGAYLDNPPGALKLTHAAKLNASLKTLEGERNQIYQISEVDKTKBELTHIKNLQ 360
DB 301 SENGAYLDNPPGALKLTHAAKLNASLKTLEGERNQIYQISEVDKTKBELTHIKNLQ 360

QY 361 TQASLOSNTFHNENQKLOKQKLVMTLYOENEMKLRKLTVEENYRLEKEEKLKSKVD 420
DB 361 TQASLOSNTFHNENQKLOKQKLVMTLYOENEMKLRKLTVEENYRLEKEEKLKSKVD 420

QY 421 EKISHATELELYRKAKDLEELERTHSYQGOIISHEKKAHDNWLAAARNLNDLR 480
DB 421 EKISHATELELYRKAKDLEELERTHSYQGOIISHEKKAHDNWLAAARNLNDLR 480

QY 481 KENAHNRQKLTETELKFELLEKDPVALDVNTAFGREHSPYGPSPPLGWPSSETRAFLSP 540
DB 481 KENAHNRQKLTETELKFELLEKDPVALDVNTAFGREHSPYGPSPPLGWPSSETRAFLSP 540

QY 541 TLLEGPLRLSPLLPGGGGRSGPGNPLDQITNERGSSCDRLTDPHRAPSDTGLSP 600
DB 541 TLLEGPLRLSPLLPGGGGRSGPGNPLDQITNERGSSCDRLTDPHRAPSDTGLSP 600

QY 601 WQDRRMFPFPGQSYDPSALPPQDRFCNSGRLSGFAELRSFNMPSLDKMDGMPSE 660
DB 601 WQDRRMFPFPGQSYDPSALPPQDRFCNSGRLSGFAELRSFNMPSLDKMDGMPSE 660

QY 661 MESSRNDTKDDLGNLVDPSSLPFAENEAATGCGFVPPPLAPITGFLFPVDARGPFLRRGPP 720
DB 661 MESSRNDTKDDLGNLVDPSSLPFAENEAATGCGFVPPPLAPITGFLFPVDARGPFLRRGPP 720

QY 721 FPPPPFGAMFGASRDYFPPRDPFGPPAPFAMENVYPPRGFPFPLPPRPGFPFPPHSEG 780
DB 721 FPPPPFGAMFGASRDYFPPRDPFGPPAPFAMENVYPPRGFPFPLPPRPGFPFPPHSEG 780

QY 781 RSEFPFGLIPPSNEPATEHPEQOET 806
DB 781 RSEFPFGLIPPSNEPATEHPEQOET 806

RESULT 2
ABG74698
ID ABG74698 standard; protein; 763 AA.
XX
AC ABG74698;
XX
DT 10-MAY-2003 (first entry)
XX

DE Human CGDD protein 7506167CD1 SEQ ID 24.
XX Human; cell growth, differentiation and death; CGDD; cardiant; cancer;
KW cytostatic; neuroprotective; gene therapy; diagnosis; cardiovascular;
XX neurological disorder.
OS Homo sapiens.
XX
XX W02003014322-A2.
XX
XX 20-FEB-2003.
XX
XX 08-AUG-2002; 2002MO-US025465.
XX
XX 08-AUG-2001; 2001US-0311017P.
XX
XX 17-AUG-2001; 2001US-0313070P.
XX
XX 17-AUG-2001; 2001US-0313071P.
XX
XX 24-AUG-2001; 2001US-0314678P.
XX
XX 31-AUG-2001; 2001US-0316892P.
XX
XX 07-SEP-2001; 2001US-0317913P.
XX
XX 14-SEP-2001; 2001US-0322182P.
XX
XX 07-DEC-2001; 2001US-0340747P.
XX
XX 20-DEC-2001; 2001US-0342761P.
XX
XX 29-MAR-2002; 2002US-0369129P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Azimzai Y, Barroso I, Baughn MR, Becha SD, Borowsky MU;
XX Duggan BM, Elliott VS, Emerling JM, Forsythe IJ, Gietzen KJ;
XX Gorvad AE, Grail RC, Griffin JA, Gururajan R, Hafalia AJA, Ison CH;
XX Kable AE, Khan FA, Lee S, Lee SY, Li JX, Reddy R, Richardson TW;
XX Sprague WW, Swarnakar A, Tang YT, Warren BA, Xu Y, Yao MG, Yue H;
XX Yue H;
XX
XX WPI; 2003-268197/26.
XX
XX N-PSDB; ABQ77437.
XX
XX New polypeptide associated with cell growth, differentiation and death,
XX useful for preparing a composition for diagnosing or treating a disease
XX e.g., cardiovascular or neurological disorders of cancer.
XX
XX Claim 79; Page 236-237; 267pp; English.
XX
XX This invention describes a novel polypeptide associated with cell growth,
XX differentiation and death (CGDD) which has cardiant, cytostatic and
XX neuroprotective activity. The polypeptides of the invention are useful
XX for preparing a composition for diagnosing, or treating a disease or
XX condition associated with decreased expression or overexpression of
XX functional CGDD e.g., cardiovascular or neurological disorders or cancer.
XX The polynucleotides of the invention can be used for gene therapy.
XX ABG74675-ABG74701 represent the human CGDD polypeptides encoded by
XX ABQ77414-ABQ77440
XX
XX Sequence 763 AA;
XX
XX Query Match 93.3%; Score 3946.5; DB 6; Length 763;
XX Best Local Similarity 94.3%; Pred. No. 2.3e-220;
XX Matches 760; Conservative 1; Mismatches 2; Indels 43; Gaps 1;
XX
XX 1 MEEPGATPOPYLGVLVEELRRVVAALPESMRDENPYGFPSELVVCAAVIGFFVLLFLW 60
XX
XX 1 MEEPGATPOPYLGVLVEELRRVVAALPESMRDENPYGFPSELVVCAAVIGFFVLLFLW 60
XX
XX 61 RSFRVSRLYVGRQKLGATLSGLIEEKCKLLEKFSLIQKEYEGYEVESSELESPEKE 120
XX
XX 61 RSFRVSRLYVGRQKLGATLSGLIEEKCKLLEKFSLIQKEYEGYEVESSELESPEKE 120
XX
XX 121 AAEARSLEATCEKLNRSNSELDEILCLEKDKQKSKHSDQDELMADISKSQISLEDE 180
XX
XX 121 AAEARSLEATCEKLNRSNSELDEILCLEKDKQKSKHSDQDELMADISKSQISLEDE 180
XX
XX 181 SKSLASQIAEAKIICTFKMSERRAIAIKDALNENSQLOTSHKOLFQEAENVKGEVSE 240
XX

Db 181 SKSLKQIAEAKICTFKMSEERRAIAIKDALNENSLQTSKQLFQQAEEVWKGEVSE 240
 Qy 241 LNKQKITFEDSKVHAEQVNDKENHKTILGHLPMKQDAVLEEDTTDDNLELVNSE 300
 Db 241 LNKQKITFEDSKVHAEQVNDKENHKTILGHLPMKQDAVLEEDTTDDNLELVNQ 300
 Qy 301 SENGAYLDNPPKALKKLIHAAKLNASLTLEGERNQIYIQLSEVDKTEELTEH:KNLQ 360
 Db 301 WENGAYLDNPPKALKKLIHAAKLNASLTLEGERNQIYIQLSEVDKTEELTEH:KNLQ 360
 Qy 361 TQASLQSENTHFENENQKLOQKLVMTLEYQENMKLHKLTVENYRLEKEEKLKVD 420
 Db 361 TQASLQSENTHFENENQKLOQKLVMTLEYQENMKLHKLTVENYRLEKEEKLKVD 420
 Qy 421 EKISHATEELETYRKRAKDLSEELERTIHSYQGIISHEKKAHDNWLAAARNAERLNLDL 480
 Db 421 EKISHATEELETYRKRAKDLSEELERTIHSYQGIISHEKKAHDNWLAAARNAERLNLDL 480
 Qy 481 KENAHNRQKLTETELKFELEKDPYALDVNTAFGREHSPYGPSP:LGWPSSETRAFLSPP 540
 Db 481 KENAHNRQKLTETELKFELEKDPYALDVNTAF----- 514
 Qy 541 TLLEGPURLSPLFGGGGRSGRFGNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSPP 600
 Db 515 -----GRSGRFGNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSPP 557
 Qy 601 WDODRRMMFPFGQSYSDSALPPQQRQDFCSNLSGRLSGPAELRSFNMPSLDKWDGSMPS 660
 Db 558 WDODRRMMFPFGQSYSDSALPPQQRQDFCSNLSGRLSGPAELRSFNMPSLDKWDGSMPS 617
 Qy 661 MESSRNDTKDDLGNLNVDPSSLSAENEAATGPGFVPPPLAPIRGFLPPVDARGPFLRRGPP 720
 Db 618 MESSRNDTKDDLGNLNVDPSSLSAENEAATGPGFVPPPLAPIRGFLPPVDARGPFLRRGPP 677
 Qy 721 FPPPPGAMFGASRDYFPDRDFGPPPPAPFAMRNVPYPPRGFPYLPYPPRGFPFPPHSE 780
 Db 678 FPPPPGAMFGASRDYFPDRDFGPPPPAPFAMRNVPYPPRGFPYLPYPPRGFPFPPHSE 737
 Qy 781 RSEFPSSGLIPPSNEPATEHPEPOQET 806
 Db 738 RSEFPSSGLIPPSNEPATEHPEPOQET 763

RESULT 3
 ABG74685
 ID ABG74685 standard; protein; 804 AA.
 AC ABG74685;
 AC ABG74685;
 DT 10-MAY-2003 (first entry)
 DE Human CGDD protein 6715627CD1 SEQ ID 11.
 DE Human CGDD protein 6715627CD1 SEQ ID 11.
 KW Human; cell growth, differentiation and death; CGDD; cardiast; cancer;
 KW cyrostatic; neuroprotective; gene therapy; diagnosis; cardiovascular;
 KW neurological disorder.
 OS Homo sapiens.
 FN WO2003014322-A2.
 PD 20-FEB-2003.
 XX 08-AUG-2002; 2002MO-US025465.
 XX 08-AUG-2001; 2001US-0311017P.
 PR 17-AUG-2001; 2001US-0313070P.
 PR 17-AUG-2001; 2001US-0313071P.
 PR 24-AUG-2001; 2001US-0314678P.
 PR 31-AUG-2001; 2001US-0316692P.
 PR 07-SEP-2001; 2001US-0317913P.
 PR 14-SEP-2001; 2001US-0322162P.
 PR 07-DEC-2001; 2001US-0340747P.

PR 20-DEC-2001; 2001US-0342761P.
 PR 29-MAR-2002; 2002US-0369129P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Azimzai Y, Barroso I, Baughn MR, Becha SD, Borowsky ML;
 PI Dugan BM, Elliott VS, Emerling BM, Forsythe IO, Gietzen KJ, Ison CH;
 PI Corvad AE, Graul RC, Griffin JA, Gururajan R, Hafalia AJA, Ison CH;
 PI Kable AE, Khan FA, Lee S, Lee SY, Li JX, Reddy R, Richardson TW;
 PI Sprague WW, Swarnakar A, Tang YT, Warren BA, Xu Y, Yao MG, Yue H;
 PI Yue H;
 XX WPI: 2003-268197/26.
 DR N-PSDB; ABQ77424.
 XX New polypeptide associated with cell growth, differentiation and death,
 PT useful for preparing a composition for diagnosing or treating a disease
 PT e.g., cardiovascular or neurological disorders or cancer.
 XX Claim 66; Page 213-215; 267pp; English.
 XX This invention describes a novel polypeptide associated with cell growth,
 CC differentiation and death (CGDD) which has cardiast, cytostatic and
 CC neuroprotective activity. The polypeptides of the invention are useful
 CC for preparing a composition for diagnosing, or treating a disease or
 CC condition associated with decreased expression or overexpression of
 CC functional CGDD e.g., cardiovascular or neurological disorders or cancer.
 CC The polynucleotides of the invention can be used for gene therapy.
 CC ABG74675-ABG74701 represent the human CGDD polypeptides encoded by
 CC ABQ77414-ABQ77440
 XX Sequence 804 AA;
 SQ

Query Match 92.1%; Score 3897; DB 6; Length 804;
 Best Local Similarity 92.7%; Fred. No. 1.8e-217; Indels 2; Gaps 2;
 Matches 747; Conservative 21; Mismatches 36;
 Qy 1 MEEPQATPQVILGLVLEELRRVVAALPESMRPDENPYGFPSELVVCAAVIGFFVVLFLW 60
 Db 1 MEEPQATPQVILGLVLEELRRVVAALPESMRPDENPYGFPSELVVCAAVIGFFVVLFLW 60
 Qy 61 RSFRSVRSRLVYGRQKLGATLSGLIEBKCLLEKFSLIQKEYEGYEVESLEDAFKE 120
 Db 61 RSFRSVRSRLVYGRQKLGATLSGLIEBKCLLEKFSLIQKEYEGYEVESLEDAFKE 120
 Qy 121 AABEARSLEATCEKLNRSNLEDEILCLEKDLQKESKHQDELMADISKSQISLEDE 180
 Db 121 AT-EAQSLEATCEKLNRSNLEDEILCLEKDLQKESKHQDELMADISKSQISLEDE 179
 Qy 181 SKSLKQIAEAKICTFKMSEERRAIAIKDALNENSLQTSKQLFQQAEEVWKGEVSE 240
 Db 180 SKSLKQIAEAKICTFKMSEERRAIAIKDALNENSLQTSKQLFQQAEEVWKGEVSE 238
 Qy 241 LNKQKITFEDSKVHAEQVNDKENHKTILGHLPMKQDAVLEEDTTDDNLELVNSE 300
 Db 239 LNKQKITFEDSKVHAEQVNDKENHKTILGHLPMKQDAVLEEDTTDDNLELVNSE 298
 Qy 301 SENGAYLDNPPKALKKLIHAAKLNASLTLEGERNQIYIQLSEVDKTEELTEH:KNLQ 360
 Db 299 SENGAYLDNPPKALKKLIHAAKLNASLTLEGERNQIYIQLSEVDKTEELTEH:KNLQ 358
 Qy 361 TQASLQSENTHFENENQKLOQKLVMTLEYQENMKLHKLTVENYRLEKEEKLKVD 420
 Db 359 TQASLQSENTHFENENQKLOQKLVMTLEYQENMKLHKLTVENYRLEKEEKLKVD 418
 Qy 421 EKISHATEELETYRKRAKDLSEELERTIHSYQGIISHEKKAHDNWLAAARNAERLNLDL 480
 Db 419 EKISHATEELETYRKRAKDLSEELERTIHSYQGIISHEKKAHDNWLAAARNAERLNLDL 478
 Qy 481 KENAHNRQKLTETELKFELEKDPYALDVNTAFGREHSPYGPSP:LGWPSSETRAFLSPP 540
 Db 479 KENAHNRQKLTETELKFELEKDPYALDVNTAFGREHSPYGPSP:LGWPSSETRAFLSPP 538

QY 541 TLLEGPLRLSPLPGGGGSGRGPNPLDHOITNERGSSCDRLTDHRAPSDTGSLSP 600
 Db 539 TLLEGPLRLSPLPGGGGSGRGPNPLDHOITNERGSSCDRLTDHRAPSDTGSLSP 598
 QY 601 WQDRRMVPPPGQSPDQALPQDRFCNSGRISGPAELRSFNMPSLDKMDGMPSE 660
 Db 599 WQDRRMVPPPGQSPDQALPQDRFCNSGRISGPAELRSFNMPSLDKMDGMPSE 658
 QY 661 MESSRNDTKDDLGNLNVDPSSLPANEATGPGFVPPPLAPIRGPIFFVDARCPFLRRGPP 720
 Db 659 MESSRNDTKDDLGNLNVDPSSLPANEATGPGFVPPPLAPIRGPIFFVDARCPFLRRGPP 718
 QY 721 FPPPPGAMFGASRDYFPPDFGPPPPAPFAMNVYPPRGFPFPLPPRGFPFPPHSE 780
 Db 719 FPPPPGAMFGASRDYFPPDFGPPPPAPFAMNVYPPRGFPFPLPPRGFPFPPHSE 778
 QY 781 RSEFPSSGLIPPSNEPATEHPEPOQET 806
 Db 779 RSEFPSSGLIPPSNEPATEHPEPOQET 804

RESULT 4
 ABO07125
 ID ABO07125 standard; protein; 806 AA.
 AC ABO07125;
 XX
 DT 13-AUG-2003 (first entry)
 XX
 DE Novel human protein NOV14a.
 XX
 KW NOVX; autoimmune disease; allergy; Alzheimer's disease; stroke;
 KW Parkinson's disease; Huntington's disease; multiple sclerosis; addiction;
 KW anxiety; pain; diabetes; glomerulonephritis; obesity;
 KW systemic lupus erythematosus; asthma; scleroderma; pancreatitis;
 KW graft versus host disease; ulcer; anaemia; cancer; trauma; infection;
 KW cardiomyopathy; atherosclerosis; hypertension; AIDS; Crohn's disease;
 KW acquired immunodeficiency syndrome; chromosomal mapping; tissue typing;
 KW forensic biology; predictive medicine; gene therapy; human.
 XX
 OS Homo sapiens.
 XX
 PN WO200298900-A2.
 XX
 PD 12-DEC-2002.
 XX
 PF 04-JUN-2002; 2002WO-US017558.
 XX
 PR 04-JUN-2001; 2001US-0295607P.
 PR 04-JUN-2001; 2001US-0295661P.
 PR 06-JUN-2001; 2001US-0296404P.
 PR 06-JUN-2001; 2001US-0296418P.
 PR 07-JUN-2001; 2001US-0296575P.
 PR 11-JUN-2001; 2001US-0297414P.
 PR 12-JUN-2001; 2001US-0297567P.
 PR 15-JUN-2001; 2001US-0298528P.
 PR 18-JUN-2001; 2001US-0299133P.
 PR 19-JUN-2001; 2001US-0299230P.
 PR 21-JUN-2001; 2001US-0299949P.
 PR 22-JUN-2001; 2001US-0300177P.
 PR 26-JUN-2001; 2001US-0300883P.
 PR 28-JUN-2001; 2001US-0301330P.
 PR 28-JUN-2001; 2001US-0301550P.
 PR 03-JUL-2001; 2001US-0302951P.
 PR 12-SEP-2001; 2001US-0318727P.
 PR 27-SEP-2001; 2001US-0325685P.
 PR 22-FEB-2002; 2002US-0358814P.
 PR 03-JUN-2002; 2002US-00161927.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Zerhusen BD, Kekuda R, Spytek KA, Shenoy SG, Miller CE, Hjal T;
 PI Gerlach VL, Baumgartner JC, Guo X, Gangolli BA, Vernet CAM;

PI Padigar M, Li L, Pena CEA, Gorman L, Anderson DW, Edinger SR;
 FI Patturajan M, Stone DJ;
 XX
 DR N-PSDS; ACD13199.
 DR
 XX
 PT Novel isolated NOVX polypeptide useful treating or preventing disorders
 PT or syndromes such as autoimmune disease, allergies, Alzheimer's disease,
 PT stroke, Parkinson's disease, Huntington's disease or multiple sclerosis.
 XX
 PS Claim 1; Page 128; 408pp; English.
 XX
 CC The invention describes an isolated NOVX polypeptide (I) comprising a
 CC sequence selected from a sequence (S1) of 1121, 635, 299, 1720, 176, 583,
 CC 214, 395, 1098, 134, 427, 1333, 407, 806, 804, 1253, 382, 1045, 284, 496,
 CC 506, 759, 390, 133, 215, 240, 1069, 116, 439, 1138, 477, 316, 269, 219,
 CC 305, 406, 460, 365, 380, 829 or 326 amino acids fully defined in the
 CC specification, and the mature form of S1. (I) is useful for treating or
 CC preventing a pathology associated with (I) in a subject, preferably
 CC human, or for identifying an agent that binds to (I), where the agent is
 CC a cellular receptor or a downstream effector. (I), a polynucleotide (II)
 CC encoding (I) or an anti-(I)-antibody (V) is useful treating or preventing
 CC disorders or syndromes such as autoimmune disease, allergies, Alzheimer's
 CC disease, stroke, Parkinson's disease, Huntington's disease, multiple
 CC sclerosis, addiction, anxiety, pain, diabetes, glomerulonephritis,
 CC systemic lupus erythematosus, asthma, scleroderma, graft versus host
 CC disease, pancreatitis, obesity, ulcers, anaemia, cancer, trauma, viral,
 CC bacterial or parasitic infections, cardiomyopathy, atherosclerosis,
 CC hypertension, acquired immunodeficiency syndrome (AIDS) or Crohn's
 CC disease. (I), (II) or (V) is useful in screening assays, detection assays
 CC (e.g., chromosomal mapping, tissue typing, forensic biology), predictive
 CC medicine (e.g., diagnostic assays, prognostic assays, monitoring clinical
 CC trials and pharmacogenomic), and in methods of treatment (e.g.,
 CC therapeutic and prophylactic). (II) is useful in gene therapy, to express
 CC (I), to detect NOVX mRNA or a genetic lesion in a NOVX gene, and to
 CC modulate NOVX activity. This is the amino acid sequence of a novel human
 CC NOV protein
 XX
 SQ Sequence 806 AA;
 Query Match 89.4%; Score 3782; DB 6; Length 806;
 Best Local Similarity 90.2%; Pred. No. 8.6e-211;
 Matches 727; Conservative 29; Mismatches 50; Indels 0; Gaps 0;
 QY 1 MEEPGATPQPYGLVLLEELARVVAALPESMRPDPENPYGFFSELVVCVAIVGFFVLLFLW 60
 Db 1 MEEPGATPQPYGLVLLEELGRVVAALPESMRPDPENPYGFFSELVVCVAIVGFFVLLFLW 60
 QY 61 RSPFSVRSRLYVGREQKLGATLSGLIEEKCKLEKFSLIQKEYEGVEVSESSLEDAFKE 120
 Db 61 RSPFSVRSRLYVGREQKLGATLSGLIEEKCKLEKFSLIQKEYEGVEVSESSLEDAFKE 120
 QY 121 AAEARSLEATCEKLNFSNSELEDEILCLEKDKQEKSKHSQOQELMADISKSISQLEDE 180
 Db 121 AAEARSLEATCEKLNFSNSELEDEILCLEKDKQEKSKHSQOQELMADISKSISQLEDE 180
 QY 181 SKLSKSIABAKIICKTFKMSERRATAIKDALNENSQLOTSQKLFQQAAYVKGQVSE 240
 Db 181 SKLSKSIABAKIICKTFKMSERRATAIKDALNENSQLOTSQKLFQQAAYVKGQVSE 240
 QY 241 LNKQKITFDSKVHAEQVINDKXNHKILTGHLPMKQQAQVLEEDTDDNLELVNSE 300
 Db 241 LNKQKITFDSKVHAEQVINDKXNHKILTGHLPMKQQAQVLEEDTDDNLELVNSE 300
 QY 301 SENGAYLDNPPKGLKKLIHAAKLNLSKLTLEGERNQIYQLSEVDKTEELTEHIKNLQ 360
 Db 301 WENGANLDDPKGALKKLIHAAKLNLSKLTLEGERNHIIQLSEVDKTEELTEHIKNLQ 360
 QY 361 TQASLOSQENHFNENKQLOOKLVKWTLEYQENEMKLRKLTVEENYRLEKEKLSKVD 420
 Db 361 TQASLOSQENYFESQKQKQKIMTFYQENEMKLYRKLTVENYRLEKEKLSRVE 420
 QY 421 EKISHATELETYRKRAKOLEEELERTHSYQGGIISHEKKAHDNWLAAARNLNDLR 480

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2004, 07:50:53 ; Search time 23 Seconds
(without alignments)
1809.154 Million cell updates/sec

Title: US-09-744-314-6

Perfect score: 4231

Sequence: 1 MEEPGATPQYLGLVLEELR.....GLPPSNEPATBHPFQOET 806

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents Ab.*

- 1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PGTUS COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	4231	100.0	806	4	US-09-976-594-321
2	270	6.4	2482	1	US-08-328-254-6
3	269	6.4	3248	1	US-08-353-700-1
4	269	6.4	3248	5	PCT-US95-16216-1
5	263	6.2	1274	4	US-09-095-443-2
6	248.5	5.9	1231	4	US-08-714-741-41
7	249.5	5.9	2954	4	US-09-150-867-1
8	249	5.9	1315	3	US-08-899-595-3
9	244.5	5.8	1581	4	US-09-866-108A-15754
10	240	5.7	1388	2	US-08-685-576-1
11	240	5.7	1388	2	US-08-685-576-4
12	240	5.7	1388	4	US-09-976-594-296
13	240	5.7	1972	4	US-08-875-435B-3
14	239.5	5.7	1354	3	US-08-688-871-2
15	239	5.6	1248	2	US-08-890-897-2
16	239	5.6	1248	3	US-09-323-735-2
17	237.5	5.6	1695	4	US-09-866-108A-15753
18	237	5.6	1972	4	US-08-875-435B-4
19	235	5.6	1886	4	US-08-938-105-3
20	234.5	5.5	1618	1	US-07-853-913-4
21	234.5	5.5	1939	4	US-09-310-187A-1
22	234.5	5.5	3878	4	US-09-914-259-11
23	233.5	5.5	1027	4	US-09-914-259-27
24	233.5	5.5	1093	5	PCT-US93-03077-1
25	232.5	5.5	816	2	US-08-533-306A-6
26	232.5	5.5	816	2	US-08-742-923A-6
27	232.5	5.5	885	2	US-08-533-306A-4

28 232.5 5.5 885 2 US-08-742-923A-4
29 231 5.5 1279 4 US-09-724-517-2
30 231 5.5 1279 4 US-09-641-807A-2
31 231 5.5 1279 4 US-09-723-096-2
32 230.5 5.4 503 4 US-09-598-287A-2
33 228 5.4 2125 4 US-09-919-172-29
34 227 5.4 976 3 US-09-104-324B-4
35 226 5.3 507 4 US-09-599-287A-24
36 225.5 5.3 631 4 US-08-477-831C-11
37 225.5 5.3 1312 2 US-08-592-126-148
38 225.5 5.3 1312 2 US-08-687-080-51
39 225.5 5.3 1312 4 US-09-158-595-148
40 224 5.3 1032 4 US-09-914-259-26
41 223 5.3 2662 4 US-09-595-684B-31
42 222.5 5.3 900 2 US-08-610-822A-62
43 222.5 5.3 900 2 US-09-005-069-62
44 222.5 5.3 900 4 US-09-171-156A-21
45 222.5 5.3 900 4 US-09-004-730A-21

ALIGNMENTS

RESULT 1
US-09-976-594-321
; Sequence 321, Application US/09976594
; Patent No. 6673549

GENERAL INFORMATION:

; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ IDS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 321
; LENGTH: 806
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 3768043CD1
US-09-976-594-321

Query Match 100.0%; Score 4231; DB 4; Length 806;
Best Local Similarity 100.0%; Pred. No. 4.9e-256; Indels 0; Gaps 0;
Matches 806; Conservative 0; Mismatches 0;
QY 1 MEEPGATPQYLGLVLEELRVRVAALPESMRPDENPYGFPSELVYCAAVIGFVLLFLW 60
|||
DB 1 MEEPGATPQYLGLVLEELRVRVAALPESMRPDENPYGFPSELVYCAAVIGFVLLFLW 60
|||
QY 61 RSPRSVRLYVGREOKLGATLSGLIECKKLEKFSLIQKEYEGYESSLEDASFKE 120
|||
DB 61 RSPRSVRLYVGREOKLGATLSGLIECKKLEKFSLIQKEYEGYESSLEDASFKE 120
|||
QY 121 AAEARSLEATCEKLNRSNSELDEILCEKDKQKSKHQDQELMADISKSIOLEDE 180
|||
DB 121 AAEARSLEATCEKLNRSNSELDEILCEKDKQKSKHQDQELMADISKSIOLEDE 180
|||
QY 181 SKSLKSOIAEAKTICTFKQSEERRAIAIKDALNENSQLOTHKQLFOQAEVWKEVSE 240
|||
DB 181 SKSLKSOIAEAKTICTFKQSEERRAIAIKDALNENSQLOTHKQLFOQAEVWKEVSE 240
|||
QY 241 LNKQKTFEEDSKVHABOVLDKENHKTITGHPMMKQAAVLEEDTTDDDDLEEVNSE 300
|||
DB 241 LNKQKTFEEDSKVHABOVLDKENHKTITGHPMMKQAAVLEEDTTDDDDLEEVNSE 300
|||
QY 301 SENGAYLDNPPKALKKLIHAAKNLASLTGEGRNQIYIQISEVDKTEBELTEHKNQ 360
|||

```

Db 301 SENGAYLDNPPKGA,KKLIHAAKLNASLKTLEGERNQIYIQLSEVDKTKTELTETHIKNLQ 360
QY 361 TQASLSQSENTHFENENOKLQKLTWMTELYDENEMKLRKLTVEENYRLKEEKLKSKVD 420
Db 361 TQASLSQSENTHFENENOKLQKLTWMTELYDENEMKLRKLTVEENYRLKEEKLKSKVD 420
QY 421 EKISHATELENYRKAADLEBELERTHSYQGOIISHEKKAHDNWLAAARNAERLNDLR 480
Db 421 EKISHATELENYRKAADLEBELERTHSYQGOIISHEKKAHDNWLAAARNAERLNDLR 480
QY 481 KENAHNRQKLTETELFELEKDPYALDVNTAFGRHSYGPSPGLGWPSSSTRFLSP 540
Db 481 KENAHNRQKLTETELFELEKDPYALDVNTAFGRHSYGPSPGLGWPSSSTRFLSP 540
QY 541 TLLEGRLRLSPLLPGGGGSGSGNPLDQITNBERGSSCDRLTDPHRAPSDTGLSP 600
Db 541 TLLEGRLRLSPLLPGGGGSGSGNPLDQITNBERGSSCDRLTDPHRAPSDTGLSP 600
QY 601 WDQDRMFPFPPQSQSYDPSALPQDRFCNSGRSLGPAELRSFNWPSLDKMDGMPSE 660
Db 601 WDQDRMFPFPPQSQSYDPSALPQDRFCNSGRSLGPAELRSFNWPSLDKMDGMPSE 660
QY 661 MESSRNDTKDDGLNVLNVPSSLPAAENATGPGVPPPLAPIRGPIFPVDARGPFLRRGPP 720
Db 661 MESSRNDTKDDGLNVLNVPSSLPAAENATGPGVPPPLAPIRGPIFPVDARGPFLRRGPP 720
QY 721 FPPPPPGAMFGASRDYFPPRDFPPGPPAPFAMENVYPPRGFPFPLPPRRPGFPPPPHSG 780
Db 721 FPPPPPGAMFGASRDYFPPRDFPPGPPAPFAMENVYPPRGFPFPLPPRRPGFPPPPHSG 780
QY 781 RSEFPGLIPPSNEPATERPEPOQET 806
Db 781 RSEFPGLIPPSNEPATERPEPOQET 806

```

```

RESULT 2
US-08-328-254-6
; Sequence 6, Application US/08328254
; Patent No. 5710022
; GENERAL INFORMATION:
; APPLICANT: Zhu, Xueliang
; APPLICANT: Lee, Wen-Hwa
; TITLE OF INVENTION: A No. 5710022el Nuclear Mitotic Phosphoprotein
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/328,254
; FILING DATE: 24-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/141,239
; FILING DATE: 22-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-CJ 1191
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2482 amino acids

```

```

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-328-254-6

```

```

Query Match 6.4%; Score 270; DB 1; Length 2482;
Best Local Similarity 21.1%; Pred. No. 2.le-08;
Matches 124; Conservative 110; Mismatches 185; Indels 170; Gaps 22;

```

```

QY 76 QKLGATLSGLIEBKCKLEKFLSIOKEYEGY-----EVESSLEDASFEKE-AAEARSLE 129
Db 1385 QTLSSDVSELKDKTKHLOEKQLSELEKDSQALSUTKCELENQIAQLNKEKELLVKESESLQ 1444
QY 130 ATC-----EKLNRN-----SLEDEILCLEKDLQKQKSKHQQDELMA 168
Db 1445 ARLSSEYKLVNSKALEAALVEKGFALRLSTQEVHQLRGIEKLRVRIEADKKQL 1504
QY 169 DISKISQSLDESKLSQIAEAKIICKTFKMEERRATAIKDALNENSLOI-----221
Db 1505 HIAEKLERERENDSLADKVENLE---RELQSEENQELVILDAENSKAEVETLKTQIEE 1561
QY 222 -----SHKQLFQOAEVWKGVSEVSNKOKITF-----ED 250
Db 1562 MARSLKVFELDLVTLRSEKENLTKIQEKQQLSELDKLLSPFKSLLEEKEQAEIQIKEE 1621
QY 251 SKVTAEQVLN-----DKE-----NHIKTLTGLHPM- 275
Db 1622 SKTAVEMQLQKELNEVAALQDQIMKATQOSLDPIIEEHQLRNSIEKLRARLEAD 1681
QY 276 MKDQAALVEEDTDD-----DNLELV---NSESENGAYLDNPPKGAUKKLIHAA 322
Db 1682 EKKQLCVLQQLKSESEHADLLKGRVENLERELELTARTNOEHALEAENSKEVETL--KA 1739
QY 323 KLNASLKTLEGERNQIYIQLSEVDKTKTELTETHIKNLQTOQASLOSENTHFN-----EN 377
Db 1740 KIEGTQSLRG-----LELDVVTTRSEKENLTNLOKECEISELEIINSSFENILQKEQ 1795
QY 378 QKLOQKLVMT--ELYQENEMKLRKLTV-----EEN-----YRLEKEBKLS 417
Db 1796 EKVQWKESSATAMEMLTQQLKELNVAALHNDQEAQKQONLSQVCELEKELQALLQ 1855
QY 418 KYDE-----KISHATELETYRKAADLEBELERTHSYQGO-----IISHEKKA 462
Db 1856 GLDEAKNNYIVLQSSVNGLIQEVEDGKQLEKDEEISRLKNQIQDOEQQLVSKLSQVEGE 1915
QY 463 HDNM-----LAARN-----AERNLNDLRKENAHRQKLTETELKFLLEKD 503
Db 1916 HOLWKEQNLRLNLTVELEQKIQLVQSKVASLQDTLEVLQSSYKYLENE 1964

```

```

RESULT 3
US-08-353-700-1
; Sequence 1, Application US/08353700
; Patent No. 5599919
; GENERAL INFORMATION:
; APPLICANT: YEN, TIMOTHY J.
; APPLICANT: RATTNER, JEROME B.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A
; TITLE OF INVENTION: TRANSIENTLY-EXPRESSED KINETOCHORE PROTEIN,
; TITLE OF INVENTION: AND METHODS OF USE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DANN, DORFMAN, HERRELL AND SKILLMAN
; STREET: 1601 MARKET STREET, SUITE 720
; CITY: PHILADELPHIA
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```

Sequence 42, Appl
Sequence 54, Appl
Sequence 56, Appl
Sequence 1117, Ap
Sequence 17, Appl
Sequence 15, Appl
Sequence 64, Appl
Sequence 16, Appl
Sequence 32312, A
Sequence 4685, Ap
Sequence 963, App
Sequence 64, Appl
Sequence 66, Appl
Sequence 106, App
Sequence 107, App
Sequence 107, App
Sequence 6751, Ap
Sequence 22285, A
Sequence 86, Appl
Sequence 2070, App
Sequence 449, App
Sequence 1586, Ap
Sequence 2, Appl
Sequence 54518, A
Sequence 7646, Ap
Sequence 838, App
Sequence 59, Appl

16 3593.5 84.9 1369 9 US-09-729-674-42
17 3190 75.4 798 15 US-10-161-493-54
18 3110 73.5 778 15 US-10-161-493-56
19 766 18.1 1193 12 US-10-296-115-1117
20 766 18.1 1884 9 US-09-785-770A-17
21 766 18.1 1884 12 US-10-409-966-17
22 766 18.1 1907 9 US-09-785-770A-16
23 766 18.1 1907 12 US-10-210-173-64
24 766 18.1 1907 12 US-10-409-966-16
25 680 16.1 136 14 US-10-029-386-32312
26 527.5 12.5 487 14 US-10-106-698-4685
27 456 10.8 352 9 US-09-925-299-963
28 456 10.8 352 10 US-09-925-299-963
29 305 7.2 1978 15 US-10-094-466-64
30 292 6.9 1591 12 US-10-210-173-66
31 276.5 6.5 1959 15 US-10-028-248A-106
32 276.5 6.5 1959 15 US-10-107-782-106
33 276 6.5 1999 15 US-10-028-248A-107
34 276 6.5 1999 15 US-10-107-782-107
35 274.5 6.5 1130 15 US-10-369-493-6751
36 271.5 6.4 1875 15 US-10-369-493-22285
37 268.5 6.3 1905 15 US-10-259-194A-86
38 268 6.3 1957 15 US-10-369-493-2070
39 267.5 6.3 1043 15 US-10-310-154-449
40 266 6.3 1790 15 US-10-369-493-1586
41 263 6.2 1274 14 US-10-020-215-2
42 263 6.2 1403 12 US-10-425-114-54518
43 263 6.2 1881 14 US-10-032-585-7646
44 262.5 6.2 860 12 US-10-072-012-836
45 262.5 6.2 860 12 US-10-037-417-59

ALIGNMENTS

RESULT 1
US-10-161-927-28
; Sequence 28, Application US/10161927
; Publication No. US20030235821A1
; GENERAL INFORMATION:
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Miller, Charles E.
; APPLICANT: Hjal, Tord
; APPLICANT: Gerlach, Valerie L.
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Vernet, Corine
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Gorman, Linda
; APPLICANT: Anderson, David W.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Patturajan, Meera
; APPLICANT: Stone, David J.
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; FILE REFERENCE: 21402-377 D (Cura 677 Other)
; CURRENT APPLICATION NUMBER: US/10/161,927
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: 60/295,661
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/295,607
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/296,404
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,418
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,575

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.
OM protein - protein search, using sw model
Run on: April 20, 2004, 07:53:18 ; Search time 47 Seconds
(without alignments)
4727.501 Million cell updates/sec

Title: US-09-744-314-6
Perfect score: 4231
Sequence: 1 MEPCATPQYGLVLELR.....GLIPSPNEPATEHPPEQET 806

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/prodata1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodata1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/prodata1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/prodata1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/prodata1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/prodata1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/prodata1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/prodata1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/prodata1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/prodata1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/prodata1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/prodata1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/prodata1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3782	89.4	806	15	US-10-161-927-28
2	3689.5	87.2	776	12	US-10-081-056-336
3	3689.5	87.2	776	12	US-10-305-654-336
4	3689.5	87.2	776	14	US-10-223-085-336
5	3689.5	87.2	776	14	US-10-223-084-336
6	3689.5	87.2	776	14	US-10-223-088-336
7	3689.5	87.2	776	14	US-10-223-090-336
8	3689.5	87.2	776	14	US-10-223-087-336
9	3689.5	87.2	776	14	US-10-223-083-336
10	3689.5	87.2	776	14	US-10-223-089-336
11	3689.5	87.2	776	14	US-10-223-081-336
12	3689.5	87.2	776	14	US-10-223-082-336
13	3634	85.9	777	9	US-09-864-761-46740
14	3626	85.7	777	9	US-09-864-761-37995
15	3595	85.0	804	15	US-10-161-927-30

;; PRIOR FILING DATE: 2001-06-07
;; PRIOR APPLICATION NUMBER: 60/297,414
;; PRIOR FILING DATE: 2001-06-11
;; PRIOR APPLICATION NUMBER: 60/297,567
;; PRIOR FILING DATE: 2001-06-12
;; PRIOR APPLICATION NUMBER: 60/298,528
;; PRIOR FILING DATE: 2001-06-15
;; PRIOR APPLICATION NUMBER: 60/325,685
;; PRIOR FILING DATE: 2001-09-27
;; PRIOR APPLICATION NUMBER: 60/299,133
;; PRIOR FILING DATE: 2001-06-18
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 190
;; SEQ ID NO 28
;; LENGTH: 806
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-161-927-28

Query Match 89.4%; Score 3782; DB 15; Length 806;
Best Local Similarity 90.2%; Pred. No. 8.5e-202;
Matches 727; Conservative 29; Mismatches 50; Indels 0; Gaps 0;

QY 1 MEEPGATQPVYGLVLELRRVVAALPESMRPDENPYGFPSELVYCAAVIGFFVLLFLW 60
DB 1 MEEPGATQPVYGLVLELGRVVAALPESMRPDENPYGFPSELVYCAAVIGFFVLLFLW 60
QY 61 RSFVRVRLVYGRQKLGATLSGLIECKLLEKFLSLIOKEYEGYEVESSELEDASTEKE 120
DB 61 RSFVRVRLVYGRQKLGATLSGLIECKLLEKFLSLIOKEYEGYEVESSELEDASTEKE 120
QY 121 AAZARSLEATCEKLNRSNSELDEILCEKDLKQEKSKSQQDELMDADISKSQSLEDE 180
DB 121 AAZARSLEATCEKLSRNSSELDEILCEKDLKQEKSKSQQDELMDADISKSQSLEDE 180
QY 181 SKSLKSQIAEAKIICKTFMSEERRAIAIKALNENSOLQTSKOLFQCEAEVVKGEYSE 240
DB 181 SKSLKSQIAEAKIICKTFMSEERRAIAIKALNENSOLQTSKOLFQCEAEVVKGEYSE 240
QY 241 LNKQKITPDSKVHAEQVLNDKENHIKTLTGHLPMKQAAVLEDDTTDDNLELVNSE 300
DB 241 LNKQKITPDSKVHAEQVLNDKENHIKTLTGHLPMKQAAVLEDDTTDDNLELVNSE 300
QY 301 SENGAYLNDPPGALKLIHAKKIASIKTLLEGERNQIYIQLSEVDKTKBELTEHKNLQ 360
DB 301 WENGANLDDPPGALKLIHAKKIASIKTLLEGERNHIIQLSEVDKTKBELTEHKNLQ 360
QY 361 TQASLQSENTHFENENQKLOQKLVMTLYQENEMKLRKLTVEENYRLEKEEKLKSV 420
DB 361 TQASLQSENIVFESENQKLOQKLVMTLYQENEMKLRKLTVEENYRLEKEEKLKSV 420
QY 421 EKISHATELYRKAQDLSEELERTHSYQOIIISHEKAKHDNLWLAARNAERNLNDR 480
DB 421 EKISHATELYRKLAKDLSEELERTVHFYQKQVISEYKRGHDNLWLAARNAERNLNDR 480
QY 481 KENAHNRQKLTETELKFELEKDPVALDVPNTAFGRHSPPYGPSPLGWPSETRAFLSP 540
DB 481 KENAHNRQKLTETELKFELEKDPVALDVPNTAFGRHSPPYGPSPLGWPSETRAFLSP 540
QY 541 TLLEGPLRLSPPLPGGGGSGSGPNPLDQHTNBERGSSCDRLTDPHRAPSDTGLSPP 600
DB 541 TLLEGPLRLSPPLPGGGGSGSGPNPLDQHTNBERGSSCDRLTDPHRAPSDTGLSPP 600
QY 601 WQDRRMFPFPGQSYPDGALPQORDCRCSNGLSGPAELRSFNMPSLDKVDSMPSE 660
DB 601 WQDRRMFPFPGQSYPDGALPQORDCRCSNGLSGPAELRSFNMPSLDKVDSMPSE 660
QY 661 MESSRNDTKDDLGNLNVPSSIPAEATGPGFVPPPLAPITGPIFPVDARGPFLRRGPP 720
DB 661 MESSRNDTKDDLGNLNVPSSIPAEATGPGFVPPPLAPITGPIFPVDARGPFLRRGPP 720
QY 721 FPPPPGAMFGASRDYFPRDRFGPPAPFAMRNIVYPRGFPYLPFRGEGFPFPPHSEG 780

Db 721 FPPPPGAMFGASRDYFPRDRFGPPAPFAMRNIVYPRGFPYLPFRGEGFPFPPHSEG 780
QY 781 RSEFPSSGLIPPSNEPATEHEPEQOET 806
Db 781 RSEFPSSGLIPPSKEPATGHPPEQOET 806

RESULT 2
US-10-081-056-336
;; Sequence 336, Application US/10081056
;; Publication No. US20040043927A1
;; GENERAL INFORMATION:
;; APPLICANT: Baker, Kevin P.
;; APPLICANT: Ferrara, Napoleone
;; APPLICANT: Gerber, Hanspeter
;; APPLICANT: Gerritsen, Mary E.
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Hillai, Kenneth J.
;; APPLICANT: Marsters, Scott A.
;; APPLICANT: Pan, James
;; APPLICANT: Paoni, Nicholas F.
;; APPLICANT: Stephan, Jean-Philippe F.
;; APPLICANT: Watanabe, Colin K.
;; APPLICANT: Wood, William I.
;; APPLICANT: Williams, P. Mickey
;; APPLICANT: Ye, Weilan
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF DISORDERS INVOLVING ANGIOGENESIS
;; FILE REFERENCE: P235PIC1
;; CURRENT APPLICATION NUMBER: US/10/081,056
;; CURRENT FILING DATE: 2002-02-20
;; PRIOR APPLICATION NUMBER: PCT/US01/21735
;; PRIOR FILING DATE: 2001-07-09
;; PRIOR APPLICATION NUMBER: US 60/219,556
;; PRIOR FILING DATE: 2000-07-20
;; PRIOR APPLICATION NUMBER: US 60/220,624
;; PRIOR FILING DATE: 2000-07-25
;; PRIOR APPLICATION NUMBER: US 60/220,664
;; PRIOR FILING DATE: 2000-07-25
;; PRIOR APPLICATION NUMBER: PCT/US00/20710
;; PRIOR FILING DATE: 2000-07-28
;; PRIOR APPLICATION NUMBER: US 60/222,695
;; PRIOR FILING DATE: 2000-08-02
;; PRIOR APPLICATION NUMBER: US 09/643,657
;; PRIOR FILING DATE: 2000-08-17
;; PRIOR APPLICATION NUMBER: PCT/US00/23522
;; PRIOR FILING DATE: 2000-08-23
;; PRIOR APPLICATION NUMBER: PCT/US00/23328
;; PRIOR FILING DATE: 2000-08-24
;; PRIOR APPLICATION NUMBER: US 60/230,978
;; PRIOR FILING DATE: 2000-09-07
;; PRIOR APPLICATION NUMBER: US 60/000,000
;; PRIOR FILING DATE: 2000-09-15
;; PRIOR APPLICATION NUMBER: US 09/664,610
;; PRIOR FILING DATE: 2000-09-18
;; PRIOR APPLICATION NUMBER: US 09/665,350
;; PRIOR FILING DATE: 2000-09-18
;; PRIOR APPLICATION NUMBER: US 60/242,922
;; PRIOR FILING DATE: 2000-10-24
;; PRIOR APPLICATION NUMBER: US 09/709,238
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: PCT/US00/30952
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: PCT/US00/30873
;; PRIOR FILING DATE: 2000-11-10
;; PRIOR APPLICATION NUMBER: PCT/US00/32678
;; PRIOR FILING DATE: 2000-12-01
;; PRIOR APPLICATION NUMBER: US 09/747,259
;; PRIOR FILING DATE: 2000-12-20
;; PRIOR APPLICATION NUMBER: PCT/US00/34956
;; PRIOR FILING DATE: 2000-12-20

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2004, 07:52:19 ; Search time 188 Seconds
(without alignments)
4184.570 Million cell updates/sec

Title: US-09-744-314-6
Perfect score: 4231
Sequence: 1 MEFGHTPPQYGLVLELR.....GLIPSPNEPATEHPPEPQET 806

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 6019581 seqs, 976053577 residues

Total number of hits satisfying chosen parameters: 6019581

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Parents AA Main:

1:	/cgn2_6/prodata/2/paa/PCTUS COMB.pcp.*	4231	100.0	806	1	PCT-US99-17167-6	Sequence 6, Appli
2:	/cgn2_6/prodata/2/paa/US06 COMB.pcp.*	4231	100.0	806	15	US-09-127-665-3	Sequence 3, Appli
3:	/cgn2_6/prodata/2/paa/US07 COMB.pcp.*	4231	100.0	806	21	US-09-744-314-6	Sequence 6, Appli
4:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	4231	100.0	806	32	US-10-753-200-321	Sequence 321, App
5:	/cgn2_6/prodata/2/paa/US09 COMB.pcp.*	4231	100.0	806	33	US-60-240-409-321	Sequence 321, App
6:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3952	93.4	862	24	US-09-948-941-369	Sequence 369, App
7:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3946.5	93.3	763	1	PCT-US02-25465-24	Sequence 24, Appli
8:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3923	92.7	804	24	US-09-948-941-336	Sequence 336, App
9:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3923	92.7	804	33	US-60-490-890-34	Sequence 34, Appli
10:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3900	92.2	804	1	PCT-US02-25465-11	Sequence 11, Appli
11:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3782	89.4	806	1	PCT-US02-17558-28	Sequence 28, Appli
12:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3782	89.4	806	27	US-10-161-927-28	Sequence 28, Appli
13:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3740	88.4	881	1	PCT-US01-08631-35639	Sequence 35639, A
14:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3712	87.7	881	1	PCT-US01-08631-50617	Sequence 50617, A
15:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3705	87.6	760	28	US-10-236-177-180	Sequence 180, App
16:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3689.5	87.2	776	26	US-10-081-056-336	Sequence 336, App
17:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3689.5	87.2	776	28	US-10-223-081-336	Sequence 336, App
18:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3689.5	87.2	776	28	US-10-223-082-336	Sequence 336, App
19:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3689.5	87.2	776	28	US-10-223-083-336	Sequence 336, App
20:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3689.5	87.2	776	28	US-10-223-084-336	Sequence 336, App
21:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3689.5	87.2	776	28	US-10-223-085-336	Sequence 336, App
22:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3689.5	87.2	776	28	US-10-223-087-336	Sequence 336, App
23:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3689.5	87.2	776	28	US-10-223-088-336	Sequence 336, App
24:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3689.5	87.2	776	28	US-10-223-089-336	Sequence 336, App
25:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3689.5	87.2	776	28	US-10-223-090-336	Sequence 336, App
26:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3674.5	86.8	776	29	US-10-305-654-336	Sequence 336, App
27:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3635	85.9	777	1	PCT-US01-08631-35638	Sequence 35638, A
28:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3634	85.9	777	1	PCT-US01-08631-50616	Sequence 50616, A
29:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3634	85.9	777	1	PCT-US01-08663-31120	Sequence 31120, A
30:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3634	85.9	777	23	US-09-864-761-46740	Sequence 46740, A
31:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3634	85.9	777	27	US-10-182-997-23198	Sequence 23198, A
32:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3634	85.9	777	28	US-10-203-135-29850	Sequence 29850, A
33:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3634	85.9	777	28	US-10-203-136-30859	Sequence 30859, A
34:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3634	85.9	777	28	US-10-203-137-31120	Sequence 31120, A
35:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3634	85.9	777	28	US-10-203-138-15126	Sequence 15126, A
36:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3634	85.9	777	28	US-10-203-138A-15126	Sequence 15126, A
37:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3634	85.9	777	28	US-10-203-139-30841	Sequence 30841, A
38:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3634	85.9	777	33	US-60-236-359-20496	Sequence 20496, A
39:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3626	85.7	777	1	PCT-US01-00663-31115	Sequence 31115, A
40:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3626	85.7	777	23	US-09-864-761-37995	Sequence 37995, A
41:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3626	85.7	777	27	US-10-182-993-30188	Sequence 30188, A
42:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3626	85.7	777	27	US-10-182-995-24467	Sequence 24467, A
43:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3626	85.7	777	27	US-10-182-997-23194	Sequence 23194, A
44:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3626	85.7	777	27	US-10-182-998-14708	Sequence 14708, A
45:	/cgn2_6/prodata/2/paa/US08 COMB.pcp.*	3626	85.7	777	28	US-10-203-135-29843	Sequence 29843, A

ALIGNMENTS

RESULT 1
PCT-US99-17167-6
; Sequence 6, Application PC/TUS9917167
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: BANDMAN, Olga
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: PATTERSON, Chandra
; APPLICANT: LAL, Preeti
; APPLICANT: BAUGHN, Mariah R.
; TITLE OF INVENTION: HUMAN CYTOSKELETAL PROTEINS
; FILE REFERENCE: PF-0568 PCT
; CURRENT APPLICATION NUMBER: PCT/US99/17167
; CURRENT FILING DATE: 1999-09-21
; PRIOR APPLICATION NUMBER: 09/127,665; unassigned; 09/128,850; unassigned
; PRIOR FILING DATE: 1998-07-31; 1998-07-31; 1998-08-04; 1998-08-04
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PERL Program
; SEQ ID NO 6

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description

; LENGTH: 806
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte Clone No: 3768043
 PCT-US99-17167-6

 Query Match 100.0%; Score 4231; DB 1; Length 806;
 Best Local Similarity 100.0%; Pred. No. 3.6e-232;
 Matches 806; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 MEEPGATPOPYLGVLVEELRRVVAALPESMRDENPYGPPSELVCAAVIGFFVLLFLW 60
 Db 1 MEEPGATPOPYLGVLVEELRRVVAALPESMRDENPYGPPSELVCAAVIGFFVLLFLW 60
 QY 61 RSFRSRLVYGRQKLGATLSGLIECKKLEKSLIQEYEGYEVESSELDASPEKE 120
 Db 61 RSFRSRLVYGRQKLGATLSGLIECKKLEKSLIQEYEGYEVESSELDASPEKE 120
 QY 121 AAEARSLEATCEKLNRSNSELDEILCLEKDKOEKSHSQOQDELMDISKSQSLDE 180
 Db 121 AAEARSLEATCEKLNRSNSELDEILCLEKDKOEKSHSQOQDELMDISKSQSLDE 180
 QY 181 SKSLASQIAEAKIICTFKMSERRAIAIKDALNENSQLOQSHKQLFOQEAENVKGEVSE 240
 Db 181 SKSLASQIAEAKIICTFKMSERRAIAIKDALNENSQLOQSHKQLFOQEAENVKGEVSE 240
 QY 241 LNKQKITFEDSKVHAEQVLDNKENHKTITGLHPMMKQAAVLEEDTDDNLELVNSE 300
 Db 241 LNKQKITFEDSKVHAEQVLDNKENHKTITGLHPMMKQAAVLEEDTDDNLELVNSE 300
 QY 301 SENGAYLDNPPKGLKLIHAAKLNASLKTLEGERNQIYIQLSEVDKTEELTEHIKNLQ 360
 Db 301 SENGAYLDNPPKGLKLIHAAKLNASLKTLEGERNQIYIQLSEVDKTEELTEHIKNLQ 360
 QY 361 TQASLSQSENTHFENENKLOQKLVMTELYQENEMKLRKLTVEENYRLKEEKLKSKVD 420
 Db 361 TQASLSQSENTHFENENKLOQKLVMTELYQENEMKLRKLTVEENYRLKEEKLKSKVD 420
 QY 421 EKISHATELETYRKAKOLEEELERTIHSYQOQIIISHEKKAHDNWLAAARNLNDLR 480
 Db 421 EKISHATELETYRKAKOLEEELERTIHSYQOQIIISHEKKAHDNWLAAARNLNDLR 480
 QY 481 KENAHNRQKLTETELKFELEKDPYALDVNTAFGREHSPYGPSPLGWSPSETRAFLSP 540
 Db 481 KENAHNRQKLTETELKFELEKDPYALDVNTAFGREHSPYGPSPLGWSPSETRAFLSP 540
 QY 541 TLLEGPLRLSPLPGGGGRSGRGNPLDHOITNERGESSCDRLTDPHRAPSDTGLSP 600
 Db 541 TLLEGPLRLSPLPGGGGRSGRGNPLDHOITNERGESSCDRLTDPHRAPSDTGLSP 600
 QY 601 WQDRRMFFPPGQSYFDSALPPQDRFCNSGRSLSGPAELRSFNMPSLDKMDGMPSE 660
 Db 601 WQDRRMFFPPGQSYFDSALPPQDRFCNSGRSLSGPAELRSFNMPSLDKMDGMPSE 660
 QY 661 MESSRNDTKDDLGNLNVDPSSLPAPNEATGFGVPPPLAPIRGFLFPVDARGFFLRGPP 720
 Db 661 MESSRNDTKDDLGNLNVDPSSLPAPNEATGFGVPPPLAPIRGFLFPVDARGFFLRGPP 720
 QY 721 FPPPPGAMFGASRDYFPPRDPFGPPAPFAMENVYPPRGGFFPYLPFRGFFPPPHSEG 780
 Db 721 FPPPPGAMFGASRDYFPPRDPFGPPAPFAMENVYPPRGGFFPYLPFRGFFPPPHSEG 780
 QY 781 RSEFFSGLIPPPNEPATEHPEQOET 806
 Db 781 RSEFFSGLIPPPNEPATEHPEQOET 806

RESULT 2
 US-09-127-665-3
 ; Sequence 3, Application US/09127665
 ; GENERAL INFORMATION:

6011518183
 Nov 2003

; APPLICANT: Bandman, Olga
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Yue, Henry
 ; APPLICANT: Corley, Neil C.
 ; APPLICANT: Guegler, Karl J.
 ; APPLICANT: Azimzai, Yalda
 ; APPLICANT: Patterson, Chandra
 ; TITLE OF INVENTION: HUMAN TROPOMYOSIN-LIKE TUMOR-ASSOCIATED PROTEINS
 ; FILE REFERENCE: PF-0568 US
 ; CURRENT APPLICATION NUMBER: US/09/127,665
 ; CURRENT FILING DATE: 1998-07-31
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 3
 ; LENGTH: 806
 ; TYPE: PRT
 ; ORGANISM: HOMO SAPIENS
 ; FEATURE:
 ; OTHER INFORMATION: 3768043, BRSTNOT24
 US-09-127-665-3

 Query Match 100.0%; Score 4231; DB 15; Length 806;
 Best Local Similarity 100.0%; Pred. No. 3.6e-232;
 Matches 806; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 MEEPGATPOPYLGVLVEELRRVVAALPESMRDENPYGPPSELVCAAVIGFFVLLFLW 60
 Db 1 MEEPGATPOPYLGVLVEELRRVVAALPESMRDENPYGPPSELVCAAVIGFFVLLFLW 60
 QY 61 RSFRSRLVYGRQKLGATLSGLIECKKLEKSLIQEYEGYEVESSELDASPEKE 120
 Db 61 RSFRSRLVYGRQKLGATLSGLIECKKLEKSLIQEYEGYEVESSELDASPEKE 120
 QY 121 AAEARSLEATCEKLNRSNSELDEILCLEKDKOEKSHSQOQDELMDISKSQSLDE 180
 Db 121 AAEARSLEATCEKLNRSNSELDEILCLEKDKOEKSHSQOQDELMDISKSQSLDE 180
 QY 181 SKSLASQIAEAKIICTFKMSERRAIAIKDALNENSQLOQSHKQLFOQEAENVKGEVSE 240
 Db 181 SKSLASQIAEAKIICTFKMSERRAIAIKDALNENSQLOQSHKQLFOQEAENVKGEVSE 240
 QY 241 LNKQKITFEDSKVHAEQVLDNKENHKTITGLHPMMKQAAVLEEDTDDNLELVNSE 300
 Db 241 LNKQKITFEDSKVHAEQVLDNKENHKTITGLHPMMKQAAVLEEDTDDNLELVNSE 300
 QY 301 SENGAYLDNPPKGLKLIHAAKLNASLKTLEGERNQIYIQLSEVDKTEELTEHIKNLQ 360
 Db 301 SENGAYLDNPPKGLKLIHAAKLNASLKTLEGERNQIYIQLSEVDKTEELTEHIKNLQ 360
 QY 361 TQASLSQSENTHFENENKLOQKLVMTELYQENEMKLRKLTVEENYRLKEEKLKSKVD 420
 Db 361 TQASLSQSENTHFENENKLOQKLVMTELYQENEMKLRKLTVEENYRLKEEKLKSKVD 420
 QY 421 EKISHATELETYRKAKOLEEELERTIHSYQOQIIISHEKKAHDNWLAAARNLNDLR 480
 Db 421 EKISHATELETYRKAKOLEEELERTIHSYQOQIIISHEKKAHDNWLAAARNLNDLR 480
 QY 481 KENAHNRQKLTETELKFELEKDPYALDVNTAFGREHSPYGPSPLGWSPSETRAFLSP 540
 Db 481 KENAHNRQKLTETELKFELEKDPYALDVNTAFGREHSPYGPSPLGWSPSETRAFLSP 540
 QY 541 TLLEGPLRLSPLPGGGGRSGRGNPLDHOITNERGESSCDRLTDPHRAPSDTGLSP 600
 Db 541 TLLEGPLRLSPLPGGGGRSGRGNPLDHOITNERGESSCDRLTDPHRAPSDTGLSP 600
 QY 601 WQDRRMFFPPGQSYFDSALPPQDRFCNSGRSLSGPAELRSFNMPSLDKMDGMPSE 660
 Db 601 WQDRRMFFPPGQSYFDSALPPQDRFCNSGRSLSGPAELRSFNMPSLDKMDGMPSE 660
 QY 661 MESSRNDTKDDLGNLNVDPSSLPAPNEATGFGVPPPLAPIRGFLFPVDARGFFLRGPP 720
 Db 661 MESSRNDTKDDLGNLNVDPSSLPAPNEATGFGVPPPLAPIRGFLFPVDARGFFLRGPP 720
 QY 721 FPPPPGAMFGASRDYFPPRDPFGPPAPFAMENVYPPRGGFFPYLPFRGFFPPPHSEG 780
 Db 721 FPPPPGAMFGASRDYFPPRDPFGPPAPFAMENVYPPRGGFFPYLPFRGFFPPPHSEG 780
 QY 781 RSEFFSGLIPPPNEPATEHPEQOET 806
 Db 781 RSEFFSGLIPPPNEPATEHPEQOET 806

QY 721 FPPPPGAMGASRDYFPPDRFPGPPAPFAMENVYPPRGFFPYLPBPPRGFFPPPPHSE 780
Db 721 FPPPPGAMGASRDYFPPDRFPGPPAPFAMENVYPPRGFFPYLPBPPRGFFPPPPHSE 780
QY 781 RSEFPGLIPPSNEPATEHPEPQET 806
Db 781 RSEFPGLIPPSNEPATEHPEPQET 806

RESULT 3

US-09-744-314-6
; Sequence 6, Application US/09744314
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: BANDMAN, Olga
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: AZIMZAI, Valda
; APPLICANT: PATTERSON, Chandra
; APPLICANT: LAL, Preeti
; APPLICANT: BAUGHN, Mariah R.
; TITLE OF INVENTION: HUMAN TROPOMYOSIN-LIKE TUMOR-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0568 PCT
; CURRENT APPLICATION NUMBER: US/09/744.314
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 09/127,665
; PRIOR FILING DATE: 1998-07-31
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq Ver. 3.0
; SEQ ID NO 6
; LENGTH: 806
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte Clone No: 3768043
US-09-744-314-6

Query Match 100.0%; Score 4231; DB 21; Length 806;
Best Local Similarity 100.0%; Pred. No. 3.6e-232;
Matches 806; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEEPGATPQYGLVLELRVVAALPESMRPDENPYGPPSELVCAAVIGFFVLLFLW 60
Db 1 MEEPGATPQYGLVLELRVVAALPESMRPDENPYGPPSELVCAAVIGFFVLLFLW 60
QY 61 RSFRSVRSRLVYGRQKLGATLSGLIECKKLEKFSLIQKEYEGYEVESSELDASFEKE 120
Db 61 RSFRSVRSRLVYGRQKLGATLSGLIECKKLEKFSLIQKEYEGYEVESSELDASFEKE 120
QY 121 AAEARSLEATCEKLNRSNSELEDEILCLEKDLKQEKSKSQDELMDADISKSIOLEDE 180
Db 121 AAEARSLEATCEKLNRSNSELEDEILCLEKDLKQEKSKSQDELMDADISKSIOLEDE 180
QY 181 SKSLKSQIAEAKIICKTFMSEERRAIAIKDALNENSQLOTSKHQKLFQOAEVWKGVESE 240
Db 181 SKSLKSQIAEAKIICKTFMSEERRAIAIKDALNENSQLOTSKHQKLFQOAEVWKGVESE 240
QY 241 LNKOKITFEDSKVHAEQVLDNKENHIKLTGHLPMKQAAVLEEDTTDDNLEVNSE 300
Db 241 LNKOKITFEDSKVHAEQVLDNKENHIKLTGHLPMKQAAVLEEDTTDDNLEVNSE 300
QY 301 SENGAYLDNPPKGAALKLIHAAKLNASLKTLEGRNQIYIQLSEVDKTEELTHIKNLQ 360
Db 301 SENGAYLDNPPKGAALKLIHAAKLNASLKTLEGRNQIYIQLSEVDKTEELTHIKNLQ 360
QY 361 TQASLQSENTHFENENKLOQKLVMTLEYQENMKLHRKLTVEENVYRLEKEEKLKVD 420
Db 361 TQASLQSENTHFENENKLOQKLVMTLEYQENMKLHRKLTVEENVYRLEKEEKLKVD 420
QY 421 EKISHATELETYRKRAKDLEELERTIHSYQGQIIISHEKKAHDNWLAAARNAERLNDLR 480

Db 421 EKISHATELETYRKRAKDLEELERTIHSYQGQIIISHEKKAHDNWLAAARNAERLNDLR 480
QY 481 KENAHNRQKLTETELKFELEKDPYALDVENTAFGREHSPYGPSPGLWMPSETAFILSP 540
Db 481 KENAHNRQKLTETELKFELEKDPYALDVENTAFGREHSPYGPSPGLWMPSETAFILSP 540
QY 541 TLLEGPIRLSLPLPGGGGSGRGPNLDHQTITNERGESSCDRLTDHRAEDSGSLSP 600
Db 541 TLLEGPIRLSLPLPGGGGSGRGPNLDHQTITNERGESSCDRLTDHRAEDSGSLSP 600
QY 601 WQDQRRMMFPPPGOSYPPDSALPPQDRFCNSGRLSGPAELRFSNMPSLDKMDGMPSE 660
Db 601 WQDQRRMMFPPPGOSYPPDSALPPQDRFCNSGRLSGPAELRFSNMPSLDKMDGMPSE 660
QY 720 MESSRNDTKDGLNLDNVDSSLPANENATGCFVPPPLAPIRGPIFPVDAGPFLRRGPP 720
Db 720 MESSRNDTKDGLNLDNVDSSLPANENATGCFVPPPLAPIRGPIFPVDAGPFLRRGPP 720
QY 780 FPPPPGAMFGASRDYFPPDRFPGPPAPFAMENVYPPRGFFPYLPBPPRGFFPPPHSE 780
Db 780 FPPPPGAMFGASRDYFPPDRFPGPPAPFAMENVYPPRGFFPYLPBPPRGFFPPPHSE 780
QY 781 RSEFPGLIPPSNEPATEHPEPQET 806
Db 781 RSEFPGLIPPSNEPATEHPEPQET 806

RESULT 4

US-10-753-200-321
; Sequence 321, Application US/10753200
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/10/753,200
; CURRENT FILING DATE: 2004-01-06
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 321
; LENGTH: 806
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 3768043CD1
US-10-753-200-321

Query Match 100.0%; Score 4231; DB 32; Length 806;
Best Local Similarity 100.0%; Pred. No. 3.6e-232;
Matches 806; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEEPGATPQYGLVLELRVVAALPESMRPDENPYGPPSELVCAAVIGFFVLLFLW 60
Db 1 MEEPGATPQYGLVLELRVVAALPESMRPDENPYGPPSELVCAAVIGFFVLLFLW 60
QY 61 RSFRSVRSRLVYGRQKLGATLSGLIECKKLEKFSLIQKEYEGYEVESSELDASFEKE 120
Db 61 RSFRSVRSRLVYGRQKLGATLSGLIECKKLEKFSLIQKEYEGYEVESSELDASFEKE 120
QY 121 AAEARSLEATCEKLNRSNSELEDEILCLEKDLKQEKSKSQDELMDADISKSIOLEDE 180
Db 121 AAEARSLEATCEKLNRSNSELEDEILCLEKDLKQEKSKSQDELMDADISKSIOLEDE 180
QY 181 SKSLKSQIAEAKIICKTFMSEERRAIAIKDALNENSQLOTSKHQKLFQOAEVWKGVESE 240
Db 181 SKSLKSQIAEAKIICKTFMSEERRAIAIKDALNENSQLOTSKHQKLFQOAEVWKGVESE 240
QY 241 LNKOKITFEDSKVHAEQVLDNKENHIKLTGHLPMKQAAVLEEDTTDDNLEVNSE 300

Handwritten signature

Db 241 LNKQKITFEDSKVHABQVNDKENHKTITGLHLPMMKQAAVLEEDTDDNLELVNSE 300
QY 301 SENGAYLDNPPKGAALKLHAAKLNASLKTLEGERNQIYIQLSEVDKTEELTEHINKLQ 360
Db 301 SENGAYLDNPPKGAALKLHAAKLNASLKTLEGERNQIYIQLSEVDKTEELTEHINKLQ 360
QY 361 TQASLOSNTHFENENKQLOKLVMTLYQENEMKLRKLTVEENYRLEKEEKLSKVD 420
Db 361 TQASLOSNTHFENENKQLOKLVMTLYQENEMKLRKLTVEENYRLEKEEKLSKVD 420
QY 421 EKISHATELETYRKAKOLEEELERTIHSYQGOIISHEKKAHDNWLAAARNAERNDLNR 480
Db 421 EKISHATELETYRKAKOLEEELERTIHSYQGOIISHEKKAHDNWLAAARNAERNDLNR 480
QY 481 KENAHNRQKLTTELKFELEKDPYALDVPNTAFGRHSPIYGPSPGLGWPSSETRAFLSPP 540
Db 481 KENAHNRQKLTTELKFELEKDPYALDVPNTAFGRHSPIYGPSPGLGWPSSETRAFLSPP 540
QY 541 TLLEGPLRLSPLPGGGGSGRGNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSPP 600
Db 541 TLLEGPLRLSPLPGGGGSGRGNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSPP 600
QY 601 WQODRRMFPFGQSYPDLSALPPQDRFCNSGRLSGPAELRSFNMPSLDKMDGSMPS 660
Db 601 WQODRRMFPFGQSYPDLSALPPQDRFCNSGRLSGPAELRSFNMPSLDKMDGSMPS 660
QY 661 MESSRNDTKDDLGNLNVDPSSLPAAENATGPGVPPPLAPIRGPIFPVDARGPFLRRGPP 720
Db 661 MESSRNDTKDDLGNLNVDPSSLPAAENATGPGVPPPLAPIRGPIFPVDARGPFLRRGPP 720
QY 721 FPPPPGAMFGASRDYFPPDRFPFPPAPFAMNVYPPRGFPFPPYLPFRGFPFPPHSE 780
Db 721 FPPPPGAMFGASRDYFPPDRFPFPPAPFAMNVYPPRGFPFPPYLPFRGFPFPPHSE 780
QY 781 RSEFPGLIPPSNEPATEHPEPOQET 806
Db 781 RSEFPGLIPPSNEPATEHPEPOQET 806

RESULT 5
US-60-240-409-321
; Sequence 321, Application US/60240409
; GENERAL INFORMATION:
; APPLICANT: Buchsinder, Jenny L.
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 P
; CURRENT APPLICATION NUMBER: US/60/240,409
; CURRENT FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 321
; LENGTH: 806
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 3768043CDI
US-60-240-409-321

Query Match 100.0%; Score 4231; DB 33; Length 806;
Best Local Similarity 100.0%; Pred. No. 3.6e-232;
Matches 806; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEEPGATPQYGLVLEELRRVVAALPESMRPDENPYGPPSELVYVCAAVIGFFVLLFLW 60
Db 1 MEEPGATPQYGLVLEELRRVVAALPESMRPDENPYGPPSELVYVCAAVIGFFVLLFLW 60
QY 61 RSRFSRSLRYVREKQKLGATLSGLIECKKLEKFSLIQKEVEGEVSESSLEDASFKE 120
Db 61 RSRFSRSLRYVREKQKLGATLSGLIECKKLEKFSLIQKEVEGEVSESSLEDASFKE 120
QY 121 AAEARSLEATCEKLNRSNSELEDEILCLEKDLQKSKHSQQDELWADISKSIOLEDE 180

Db 121 AAEARSLEATCEKLNRSNSELEDEILCLEKDLQKSKHSQQDELWADISKSIOLEDE 180
QY 181 SKLSKQIAEAKIICKTFKMSERRAIAIKDALNENSQLOTSKQLFQOEAEVWKGEVSE 240
Db 181 SKLSKQIAEAKIICKTFKMSERRAIAIKDALNENSQLOTSKQLFQOEAEVWKGEVSE 240
QY 241 LNKQKITFEDSKVHABQVNDKENHKTITGLHLPMMKQAAVLEEDTDDNLELVNSE 300
Db 241 LNKQKITFEDSKVHABQVNDKENHKTITGLHLPMMKQAAVLEEDTDDNLELVNSE 300
QY 301 SENGAYLDNPPKGAALKLHAAKLNASLKTLEGERNQIYIQLSEVDKTEELTEHINKLQ 360
Db 301 SENGAYLDNPPKGAALKLHAAKLNASLKTLEGERNQIYIQLSEVDKTEELTEHINKLQ 360
QY 361 TQASLOSNTHFENENKQLOKLVMTLYQENEMKLRKLTVEENYRLEKEEKLSKVD 420
Db 361 TQASLOSNTHFENENKQLOKLVMTLYQENEMKLRKLTVEENYRLEKEEKLSKVD 420
QY 421 EKISHATELETYRKAKOLEEELERTIHSYQGOIISHEKKAHDNWLAAARNAERNDLNR 480
Db 421 EKISHATELETYRKAKOLEEELERTIHSYQGOIISHEKKAHDNWLAAARNAERNDLNR 480
QY 481 KENAHNRQKLTTELKFELEKDPYALDVPNTAFGRHSPIYGPSPGLGWPSSETRAFLSPP 540
Db 481 KENAHNRQKLTTELKFELEKDPYALDVPNTAFGRHSPIYGPSPGLGWPSSETRAFLSPP 540
QY 541 TLLEGPLRLSPLPGGGGSGRGNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSPP 600
Db 541 TLLEGPLRLSPLPGGGGSGRGNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSPP 600
QY 601 WQODRRMFPFGQSYPDLSALPPQDRFCNSGRLSGPAELRSFNMPSLDKMDGSMPS 660
Db 601 WQODRRMFPFGQSYPDLSALPPQDRFCNSGRLSGPAELRSFNMPSLDKMDGSMPS 660
QY 661 MESSRNDTKDDLGNLNVDPSSLPAAENATGPGVPPPLAPIRGPIFPVDARGPFLRRGPP 720
Db 661 MESSRNDTKDDLGNLNVDPSSLPAAENATGPGVPPPLAPIRGPIFPVDARGPFLRRGPP 720
QY 721 FPPPPGAMFGASRDYFPPDRFPFPPAPFAMNVYPPRGFPFPPYLPFRGFPFPPHSE 780
Db 721 FPPPPGAMFGASRDYFPPDRFPFPPAPFAMNVYPPRGFPFPPYLPFRGFPFPPHSE 780
QY 781 RSEFPGLIPPSNEPATEHPEPOQET 806
Db 781 RSEFPGLIPPSNEPATEHPEPOQET 806

RESULT 6
US-09-948-941-369
; Sequence 369, Application US/09948941
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO00788
; CURRENT APPLICATION NUMBER: US/09/948,941
; CURRENT FILING DATE: 2001-09-10
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 12618
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 369
; LENGTH: 862
; TYPE: PRT
; ORGANISM: Human
US-09-948-941-369
Query Match 93.4%; Score 3952; DB 24; Length 862;
Best Local Similarity 93.7%; Pred. No. 3.2e-216;
Matches 755; Conservative 20; Mismatches 29; Indels 2; Gaps 2;
QY 1 MEEPGATPQYGLVLEELRRVVAALPESMRPDENPYGPPSELVYVCAAVIGFFVLLFLW 60

Db 59 MEEFGATPQPLGLLLELRVVAALPFGMEPDSNLYGFPWELVICAUVGVFAVLLFLW 118
Qy 61 RSFRSVRSRLVYVGRQKLGATLSGLIEBKCKLLEKFSLIQKEYEGYEVSSLEDASFEKE 120
Db 119 RSFRSVRSRLVYVGRQKLGATLSGLIEBKCKLLEKFSLIQKEYEGYEVSSLEDASFEKE 178
Qy 121 AAEARSLEATCEKLNRSNLEDEILCLEKDLKQEKSKHSDQDELMDADISKSIOLEDE 180
Db 179 AT-EAQSLEATCEKLNRSNLEDEILCLEKDLKQEKSKHSDQDELMDADISKSIOLEDE 237
Qy 181 SKSLKSQIAEAKIICKTFKMSERRAIAIKDALNENSQLOTSKHQKLFQQAEEVWKGVESE 240
Db 238 SKSLKSQIAEAKIICKTFKMSERRAIAIKDALNENSQLOTSKHQKLFQQAEEVWKGVESE 296
Qy 241 LNKOKITFEDSKVHAEQVLDKNDKHEIKTLTGHLPMDKQAAVLDEEDTTDDNLELVNSE 300
Db 297 LNKOKITFEDSKVHAEQVLDKNDKHEIKTLTGHLPMDKQAAVLDEEDTTDDNLELVNSE 356
Qy 301 SENGAYLDNPPKGLKLIHAAKLNASLKTLEGRNQIYIQLSEVDKTKBELTEHINKLQ 360
Db 357 SENGAYLDNPPKGLKLIHAAKLNASLKTLEGRNQIYIQLSEVDKTKBELTEHINKLQ 416
Qy 361 TQASLASQENHFNENKQLOKLVMTLYQENEMKLRKLTVEENYRLEKEEKSKVD 420
Db 417 TQASLASQENHFNENKQLOKLVMTLYQENEMKLRKLTVEENYRLEKEEKSKVD 476
Qy 421 EKISHATEETLYRKRAKDLEELERTIHSYQGGIISHEKKAHNDWLAARNAERNLNDR 480
Db 477 EKISHATEETLYRKRAKDLEELERTIHSYQGGIISHEKKAHNDWLAARNAERNLNDR 536
Qy 481 KENAHNRQKLTETELKELLEKDYALDVNTAFGRHSYGRSPGLWSPSETRAFUSPP 540
Db 537 KENAHNRQKLTETELKELLEKDYALDVNTAFGRHSYGRSPGLWSPSETRAFUSPP 596
Qy 541 TLLEGPLRLSPLPGLGGGSRGFGNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSPP 600
Db 597 TLLEGPLRLSPLPGLGGGSRGFGNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSPP 656
Qy 601 WDQRRMFPFGQSYDPSALPPQDRFCNSGRSLSPALERSFNPSLDKVDGSMSE 660
Db 657 WDQRRMFPFGQSYDPSALPPQDRFCNSGRSLSPALERSFNPSLDKVDGSMSE 716
Qy 661 MESSRNDTKDDLGNLVNPDSSLPENATGFGFVPPPLAPIRGFLFPVDARGPFLRGPP 720
Db 717 MESSRNDTKDDLGNLVNPDSSLPENATGFGFVPPPLAPIRGFLFPVDARGPFLRGPP 776
Qy 721 FPPPPPGAMFGASRDYFPPRPPGPPAPFAMRVNVPFRGPPVLPFRPPGPPPHSEG 780
Db 777 FPPPPPGAMFGASRDYFPPRPPGPPAPFAMRVNVPFRGPPVLPFRPPGPPPHSEG 836
Qy 781 RSEFPGLIPPSNBPATHEPPOQET 806
Db 837 RSEFPGLIPPSNBPATHEPPOQET 862

RESULT 7
PCT-US02-25465-24

; Sequence 24, Application PC/ITUS0225465
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: AZIMZAL, Yalda
; APPLICANT: BARROSO, Ines
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: BECHA, Shanya D.
; APPLICANT: BOROWSKY, Mark L.
; APPLICANT: DUGGAN, Brendan M.
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: EMERLING, Brooke M.
; APPLICANT: FORTSYTHE, Ian J.
; APPLICANT: GIETZEN, Kimberly J.
; APPLICANT: GORVAD, Ann E.
; APPLICANT: GRAUL, Richard C.

; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: GURURAJAN, Rajagopal
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: ISON, Craig H.
; APPLICANT: KABLE, Amy E.
; APPLICANT: KAHN, Farrah A.
; APPLICANT: LEE, Sally
; APPLICANT: LEE, Soo Yeun
; APPLICANT: LI, Joana X.
; APPLICANT: REDDY, Roopa
; APPLICANT: RICHARDSON, Thomas W.
; APPLICANT: SPRAGUE, William W.
; APPLICANT: SWARNAKAR, Anita
; APPLICANT: TANG, Y. Tom
; APPLICANT: WARREN, Bridget A.
; APPLICANT: XU, Yuming
; APPLICANT: YAO, Monique G.
; APPLICANT: YUE, Henry
; APPLICANT: YUE, Huibin
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PF-1126 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/25465
; CURRENT FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: US 60/311,017
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/313,070
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/313,071
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/314,678
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/316,692
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/317,913
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/322,182
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/340,747
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/342,761
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/369,129
; PRIOR FILING DATE: 2002-03-29
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PERL Program
; SEQ ID NO 24
; LENGTH: 763
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Misc feature
; OTHER INFORMATION: Incyte ID No: 7506167CD1
; PCT-US02-25465-24

Query Match 93.3%; Score 3946.5; DB 1; Length 763;
Best Local Similarity 94.3%; Pred. No. 5.6e-216;
Matches 760; Conservative 1; Mismatches 2; Indels 43; Gaps 1;
Qy 1 MEEFGATPQPLGLLLELRVVAALPESMRPDENPYGPPSELVCAAVIGFVLLFLW 60
Db 1 MEEFGATPQPLGLLLELRVVAALPESMRPDENPYGPPSELVCAAVIGFVLLFLW 60
Qy 61 RSFRSVRSRLVYVGRQKLGATLSGLIEBKCKLLEKFSLIQKEYEGYEVSSLEDASFEKE 120
Db 61 RSFRSVRSRLVYVGRQKLGATLSGLIEBKCKLLEKFSLIQKEYEGYEVSSLEDASFEKE 120
Qy 121 AAEARSLEATCEKLNRSNLEDEILCLEKDLKQEKSKHSDQDELMDADISKSIOLEDE 180
Db 121 AAEARSLEATCEKLNRSNLEDEILCLEKDLKQEKSKHSDQDELMDADISKSIOLEDE 180
Qy 181 SKSLKSQIAEAKIICKTFKMSERRAIAIKDALNENSQLOTSKHQKLFQQAEEVWKGVESE 240
Db 181 SKSLKSQIAEAKIICKTFKMSERRAIAIKDALNENSQLOTSKHQKLFQQAEEVWKGVESE 240

QY 241 LNKQKITPDSKVHAEQVNDKXENIKTLTGHLPMKQDQAAVLEDDTTDDNDNLELVNSE 300
 Db 241 LNKQKITPDSKVHAEQVNDKXENIKTLTGHLPMKQDQAAVLEDDTTDDNDNLELVNSQ 300
 QY 301 SENGAYLDNPPKGAALKKLIHAAKLNASLKTLEGERNQIYIQLSEVDKTKKEELTEHINKLQ 360
 Db 301 WENGAYLDNPPKGAALKKLIHAAKLNASLKTLEGERNQIYIQLSEVDKTKKEELTEHINKLQ 360
 QY 361 TQASLOSNTHEFNENQKLOQKLVMTELYQENEMKLRKLTVEENYRLSEKEELSKVD 420
 Db 361 TQASLOSNTHEFNENQKLOQKLVMTELYQENEMKLRKLTVEENYRLSEKEELSKVD 420
 QY 421 EKISHATELETYRKRAKOLEELEERTIHSYQGOIISHEKKAHDNWLAAARNAERNLNDLR 480
 Db 421 EKISHATELETYRKRAKOLEELEERTIHSYQGOIISHEKKAHDNWLAAARNAERNLNDLR 480
 QY 481 KENAHNRQKLTETELKPELLEKDPYALDVPNTAFGREHSPYGPSPGLGWPSSSETRAFLSPP 540
 Db 481 KENAHNRQKLTETELKPELLEKDPYALDVPNTAFGREHSPYGPSPGLGWPSSSETRAFLSPP 540
 QY 541 TLLEGPLRLSPLPGGGGSGSGPGNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSPP 600
 Db 541 TLLEGPLRLSPLPGGGGSGSGPGNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSPP 600
 QY 515 -----GRSGSGPGNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSPP 557
 Db 515 -----GRSGSGPGNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSPP 557
 QY 601 WDQDRMMFPFGQSYDPSALPPQDRFCNSGRLSGPAELRSFNMPSLDKMDGSMPS 660
 Db 558 WDQDRMMFPFGQSYDPSALPPQDRFCNSGRLSGPAELRSFNMPSLDKMDGSMPS 617
 QY 661 MESSRNDTKDDLGNLNVDPSSLPAAENATGPGFVPPPLAPIRGPLFPVDARGPFLRRGPP 720
 Db 618 MESSRNDTKDDLGNLNVDPSSLPAAENATGPGFVPPPLAPIRGPLFPVDARGPFLRRGPP 677
 QY 721 FPPPPGAMFGASRDYFPPDRFPFPAPFAMRVNVPYPRGFPYLPFRPGFPFPFPSPHSEG 780
 Db 678 FPPPPGAMFGASRDYFPPDRFPFPAPFAMRVNVPYPRGFPYLPFRPGFPFPFPSPHSEG 737
 QY 781 RSEFPGSLIPPSNEPATEHPEPOQET 806
 Db 738 RSEFPGSLIPPSNEPATEHPEPOQET 763

RESULT 8
 US-09-948-941-336
 ; Sequence 336, Application US/09948941
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; FILE REFERENCE: WITH CANCER, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CLO000788
 ; CURRENT APPLICATION NUMBER: US/09/948,941
 ; PRIOR FILING DATE: 2001-09-10
 ; PRIOR APPLICATION NUMBER: 60/231,328
 ; PRIOR FILING DATE:
 ; NUMBER OF SEQ ID NOS: 12618
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 336
 ; LENGTH: 804
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-948-941-336

Query Match 92.7%; Score 3923; DB 24; Length 804;
 Best Local Similarity 93.1%; Pred. No. 1.3e-214;
 Matches 750; Conservative 21; Mismatches 33; Indels 2; Gaps 2;

QY 1 MEEPGATPQYLGLVLEELRRVVAALPESMRDENPYGPSELVYCAAVIGFVLLFLW 60
 Db 1 MEEPGATPQYLGLVLEELRRVVAALPESMRDENPYGPSELVYCAAVIGFVLLFLW 60
 QY 61 RSRFSRSLVYRGKLGATISGLIEEKCKLEKFLSLQKEVGEYVESSLEDASFEKE 120
 Db 61 RSRFSRSLVYRGKLGATISGLIEEKCKLEKFLSLQKEVGEYVESSLEDASFEKE 120

QY 121 AAEASASLEATCEKLNRSNSELEDEILCLEKDLKQEKSKHSGQODELMADISKSIOSEDE 180
 Db 121 AT-EAOSLEATCEKLNRSNSELEDEILCLEKELKEEKSKHSGQODELMADISKSIOSEDE 179
 QY 181 SKSLKQIAEAKIICITFXMSBERRAIAIKDALNENSQLOTSHKQLFQQAEEVWKEVSE 240
 Db 180 SKSLKQIAEAKIICITFXMSBERRAIAIKDALNENSQLOTSHKQLFQQAEEVWKEVSE 238
 QY 241 LNKQKITPDSKVHAEQVNDKXENIKTLTGHLPMKQDQAAVLEDDTTDDNDNLELVNSE 300
 Db 239 LNKQKITPDSKVHAEQVNDKXENIKTLTGHLPMKQDQAAVLEDDTTDDNDNLELVNSE 298
 QY 301 SENGAYLDNPPKGAALKKLIHAAKLNASLKTLEGERNQIYIQLSEVDKTKKEELTEHINKLQ 360
 Db 299 SENGAYLDNPPKGAALKKLIHAAKLNASLKTLEGERNQIYIQLSEVDKTKKEELTEHINKLQ 358
 QY 361 TQASLOSNTHEFNENQKLOQKLVMTELYQENEMKLRKLTVEENYRLSEKEELSKVD 420
 Db 359 TQASLOSNTHEFNENQKLOQKLVMTELYQENEMKLRKLTVEENYRLSEKEELSKVD 418
 QY 421 EKISHATELETYRKRAKOLEELEERTIHSYQGOIISHEKKAHDNWLAAARNAERNLNDLR 480
 Db 419 EKISHATELETYRKRAKOLEELEERTIHSYQGOIISHEKKAHDNWLAAARNAERNLNDLR 478
 QY 481 KENAHNRQKLTETELKPELLEKDPYALDVPNTAFGREHSPYGPSPGLGWPSSSETRAFLSPP 540
 Db 479 KENAHNRQKLTETELKPELLEKDPYALDVPNTAFGREHSPYGPSPGLGWPSSSETRAFLSPP 538
 QY 541 TLLEGPLRLSPLPGGGGSGSGPGNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSPP 600
 Db 539 TLLEGPLRLSPLPGGGGSGSGPGNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSPP 598
 QY 601 WDQDRMMFPFGQSYDPSALPPQDRFCNSGRLSGPAELRSFNMPSLDKMDGSMPS 660
 Db 599 WDQDRMMFPFGQSYDPSALPPQDRFCNSGRLSGPAELRSFNMPSLDKMDGSMPS 658
 QY 661 MESSRNDTKDDLGNLNVDPSSLPAAENATGPGFVPPPLAPIRGPLFPVDARGPFLRRGPP 720
 Db 659 MESSRNDTKDDLGNLNVDPSSLPAAENATGPGFVPPPLAPIRGPLFPVDARGPFLRRGPP 718
 QY 721 FPPPPGAMFGASRDYFPPDRFPFPAPFAMRVNVPYPRGFPYLPFRPGFPFPFPSPHSEG 780
 Db 719 FPPPPGAMFGASRDYFPPDRFPFPAPFAMRVNVPYPRGFPYLPFRPGFPFPFPSPHSEG 778
 QY 781 RSEFPGSLIPPSNEPATEHPEPOQET 806
 Db 779 RSEFPGSLIPPSNEPATEHPEPOQET 804

RESULT 9
 US-60-490-890-34
 ; Sequence 34, Application US/60490890
 ; GENERAL INFORMATION:
 ; APPLICANT: Li, Martha
 ; APPLICANT: Rupnow, Brent A.
 ; APPLICANT: Webster, Kevin R.
 ; APPLICANT: Jackson, Donald
 ; APPLICANT: Wong, Tai W.
 ; TITLE OF INVENTION: BIOMARKERS OF CYCLIN-DEPENDENT KINASE MODULATION
 ; FILE REFERENCE: D0310 PSP
 ; CURRENT APPLICATION NUMBER: US/60/490,890
 ; CURRENT FILING DATE: 2003-07-29
 ; NUMBER OF SEQ ID NOS: 2779
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 34
 ; LENGTH: 804
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-60-490-890-34

Query Match 92.7%; Score 3923; DB 33; Length 804;
 Best Local Similarity 93.1%; Pred. No. 1.3e-214;

```
Matches 750; Conservative 21; Mismatches 33; Indels 2; Gaps 2;
QY 1 MEEPGATPQYGLVLEELRRVVAALPESMRDENPYGFPSELVYCAAVIGFVLLFLW 60
Db 1 MEEPGATPQYGLVLEELRRVVAALPEGMRPDSNLYGFPWELVICAAGVGFVLLFLW 60
QY 61 RSFRSVRSRLVYVREKQKGLATLSGLIEEKCKLEKFLSLQKEVGEYEVESSELDASFEKE 120
Db 61 RSFRSVRSRLVYVREKQKGLATLSGLIEEKCKLEKFLSLQKEVGEYEVESSELDASFEKE 120
QY 121 AAEARSLEATCEKLNRSNSELDEILCLEKQKSHSQDDELADISKSIOLEDE 180
Db 121 AT-EAQSLEATCEKLNRSNSELDEILCLEKQKSHSQDDELADISKSIOLEDE 179
QY 181 SKLSKQIAEAKIICKTFKMSERRAIAIKDALNENSQLOTSKQLPQOEAEVWKEVSE 240
Db 180 SKLSKQIAEAKTFOIPNNEERLATAIKDALNENSQLOESQKQLL-QEAEVWKEVSE 238
QY 241 LNKQKITFEDSKVHAEQVLDNKENHKTGLTHLPMMKQQAUVLEEDTDDNLELEVNSE 300
Db 239 LNKQKITFEDSKVHAEQVLDNKENHKTGLTHLPMMKQQAUVLEEDTDDNLELEVNSE 298
QY 301 SENGAYLDNPKGALKLIHAAKLNASLKTLEGERNOIYIQLSEVDKTEELTEHKNLQ 360
Db 299 SENGAYLDNPKGALKLIHAAKLNASLKTLEGERNOIYIQLSEVDKTEELTEHKNLQ 358
QY 361 TQASLQSENTHFENENQKLOQLKVMTELYQENEMKLRKLTVEENYRLEKEKLSKVD 420
Db 359 TQASLQSENTHFENENQKLOQLKVMTELYQENEMKLRKLTVEENYRLEKEKLSKVD 418
QY 421 EKISHATEELETYRKRAKDLLEELERTIHSYQGOIISHEKKAHNDMLAARNLNDLR 480
Db 419 EKISHATEELETYRKRAKDLLEELERTIHSYQGOIISHEKKAHNDMLAARNLNDLR 478
QY 481 KENAHNQKLTETELKELLEKOPYALDVNTAFGRHSPYGPSPGLWPSSETFLSP 540
Db 479 KENAHNQKLTETELKELLEKOPYALDVNTAFGRHSPYGPSPGLWPSSETFLSP 538
QY 541 TLLEGPLRLSPLPGGGRSGRSGNPLDHQITNERGESSCDRLTDPHRAPSDTGSLSP 600
Db 539 TLLEGPLRLSPLPGGGRSGRSGNPLDHQITNERGESSCDRLTDPHRAPSDTGSLSP 598
QY 601 WDQRRMMFPFGQSYFDSALPPQORQFCSNSGRLSGPAELRSFNMFSLDKVDGSMPS 660
Db 599 WDQRRMMFPFGQSYFDSALPPQORQFCSNSGRLSGPAELRSFNMFSLDKVDGSMPS 658
QY 661 MESSRNDTKDGLNLPDSSLPAENATGPGFVPPPLAIPRGFLFPVDARGPFLRRGPP 720
Db 659 MESSRNDTKDGLNLPDSSLPAENATGPGFVPPPLAIPRGFLFPVDARGPFLRRGPP 718
QY 721 FPPPPGAMFGASRDYFPPRDFGPPAPPAMRNVPYPRGFPYLPYPRGFPFPPHSE 780
Db 719 FPPPPGAMFGASRDYFPPRDFGPPAPPAMRNVPYPRGFPYLPYPRGFPFPPHSE 778
QY 781 RSEFPGLIPPSNPATEHPPEQET 806
Db 779 RSEFPGLIPPSNPATEHPPEQET 804

RESULT 10
PCT-US02-25465-11
; Sequence 11, Application PC/TUS0225465
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BARROSO, Ines
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: BECHA, Shanya D.
; APPLICANT: BOROWSKI, Mark L.
; APPLICANT: DUGGAN, Brendan M.
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: EMERLING, Brooke M.
; APPLICANT: FORSYTHE, Ian J.
```

```
; APPLICANT: GIETZEN, Kimberly J.
; APPLICANT: GORVAD, Ann E.
; APPLICANT: GRAUL, Richard C.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: GURURAJAN, Rajagopal
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: ISON, Craig H.
; APPLICANT: KABLE, Amy E.
; APPLICANT: KHAN, Farrah A.
; APPLICANT: LEE, Sally
; APPLICANT: LEE, Soo Yeun
; APPLICANT: LI, Joana X.
; APPLICANT: REDDY, Roopa
; APPLICANT: RICHARDON, Thomas W.
; APPLICANT: SPRAGUE, William W.
; APPLICANT: SWARNAKAR, Anita
; APPLICANT: TANG, Y. Tom
; APPLICANT: WARREN, Bridget A.
; APPLICANT: XU, Yuming
; APPLICANT: YAO, Monique G.
; APPLICANT: YUE, Henry
; APPLICANT: YUE, Huibin
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PF-1126 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/25465
; CURRENT FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: US 60/311,017
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/313,070
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/313,071
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/314,678
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/316,692
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/317,913
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/322,182
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/340,747
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/342,761
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/369,129
; PRIOR FILING DATE: 2002-03-29
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PERL Program
; SEQ ID NO 11
; LENGTH: 804
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 6715627CD1
PCT-US02-25465-11

Query Match 92.2%; Score 3900; DB 1; Length 804;
Best Local Similarity 92.8%; Pred. No. 2.7e-213;
Matches 748; Conservative 20; Mismatches 36; Indels 2; Gaps 2;
QY 1 MEEPGATPQYGLVLEELRRVVAALPESMRDENPYGFPSELVYCAAVIGFVLLFLW 60
Db 1 MEEPGATPQYGLVLEELRRVVAALPEGMRPDSNLYGFPWELVICAAGVGFVLLFLW 60
QY 61 RSFRSVRSRLVYVREKQKGLATLSGLIEEKCKLEKFLSLQKEVGEYEVESSELDASFEKE 120
Db 61 RSFRSVRSRLVYVREKQKGLATLSGLIEEKCKLEKFLSLQKEVGEYEVESSELDASFEKE 120
QY 121 AAEARSLEATCEKLNRSNSELDEILCLEKQKSHSQDDELADISKSIOLEDE 180
Db 121 AT-EAQSLEATCEKLNRSNSELDEILCLEKQKSHSQDDELADISKSIOLEDE 179
```


361 TQASLOSINTHENENOKLOKLUKUMTELYQENEMKLRKLTVEENYALEKEEKLKVD 420
Db TQASLOSINTHENENOKLOKLUKUMTELYQENEMKLRKLTVEENYALEKEEKLKVD 420
421 EKISHATEELETYRKAKOLEEELERTIHSYQOQIISHEKKAHDNWLAAARNALNDLR 480
Db EKISHATEELETYRKAKOLEEELERTIHSYQOQIISHEKKAHDNWLAAARNALNDLR 480
481 KENAHNRQKLTETELKELLEKDPYALDVNTAFGREHSPYSGPLGWSSETRAFLSP 540
Db KENAHNRQKLTETELKELLEKDPYALDVNTAFGREHSPYSGPLGWSSETRAFLSP 540
541 TLLEGLRLSLPLPGGGGRSGRPGNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSP 600
Db TLLEGLRLSLPLPGGGGRSGRPGNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSP 600
601 WQDRRMFPPPGQSYDPSALPORQDRFCNSGRSLSGPAELRSFNMPSLDKMDGSMSE 660
Db WQDRRMFPPPGQSYDPSALPORQDRFCNSGRSLSGPAELRSFNMPSLDKMDGSMSE 660
661 MESSRNDTKDDLGNLNVDSLSLPAENATGPGFVPPPLAIPRGPLFPVDARGPFLRRGPP 720
Db MESSRNDTKDDLGNLNVDSLSLPAENATGPGFVPPPLAIPRGPLFPVDARGPFLRRGPP 720
721 FPPPPGAMFGASRDYFPFRDFFGPPAPFAMRNVPYPRGFPYLPFRGFPFPPHSEG 780
Db FPPPPGAMFGASRDYFPFRDFFGPPAPFAMRNVPYPRGFPYLPFRGFPFPPHSEG 780
781 RSEFPGSLIPPSNEPATEHEPQOET 806
Db RSEFPGSLIPPSNEPATEHEPQOET 806

RESULT 12
US-10-161-927-28
; Sequence 28 Application US/10161927
; GENERAL INFORMATION:
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spylek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Miller, Charles E.
; APPLICANT: Hjal, Tord
; APPLICANT: Gerlach, Valerie L.
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Vernet, Corine
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Gorman, Linda
; APPLICANT: Anderson, David W.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Patturajan, Meera
; APPLICANT: Stone, David J.
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS
; FILE REFERENCE: 21402-377 D (Cura 677 Other)
; CURRENT APPLICATION NUMBER: US/10/161.927
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: 60/295,661
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/295,607
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/296,404
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,418
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,575
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/297,414
; PRIOR FILING DATE: 2001-06-11

PRIOR APPLICATION NUMBER: 60/297,567
PRIOR FILING DATE: 2001-06-12
PRIOR APPLICATION NUMBER: 60/298,528
PRIOR FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: 60/325,685
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: 60/299,133
PRIOR FILING DATE: 2001-06-18
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 190
SEQ ID NO 28
LENGTH: 806
TYPE: PRT
ORGANISM: Homo sapiens
US-10-161-927-28

Query Match 89.4%; Score 3782; DB 27; Length 806;
Best Local Similarity 90.2%; Pred. No. 1.5e-206; Indels 0; Gaps 0;
Matches 727; Conservative 29; Mismatches 50;

QY 1 MEEPGATQPVLGLVLEELRRVVAALPESMRPDENPYGFPSELVVCAAVIGFVVLLFLW 60
Db 1 MEEPGATQPVLGLVLEELRRVVAALPESMRPDENPYGFPSELVVCAAVIGFVVLLFLW 60
61 RSFRSVRLVYGRQKLGATLSGLIEBKCKLEKFSLIQKEYGYEVESSELDASPEKE 120
Db 61 RSFRSVRLVYGRQKLGATLSGLIEBKCKLEKFSLIQKEYGYEVESSELDASPEKE 120
121 AAEARSLEATCEKLNRSNSELDEILCLEKDLKQEKSKHQDQELMADISKSIOSELE 180
Db 121 AAEARSLEATCEKLNRSNSELDEILCLEKDLKQEKSKHQDQELMADISKSIOSELE 180
181 SKLSQSQTAEAKIIKCTFMSEERRAIAIKALNENSOLOTSKQOLFOEAEVWKEVSE 240
Db 181 SKLSQSQTAEAKIIKCTFMSEERRAIAIKALNENSOLOTSKQOLFOEAEVWKEVSE 240
241 LNKQKITFEDSKVHAEQVNLNDKENHIKTLTGHLPWKQDAVLEEDTDDNLELVNSE 300
Db 241 LNKQKITFEDSKVHAEQVNLNDKENHIKTLTGHLPWKQDAVLEEDTDDNLELVNSE 300
301 SENGAYLDNPPKALKLIHAALNASLKTJEGERNQYIQLSEVDKTEELTEHILKQL 360
Db 301 WENGANLDDPPKALKLIHAALNVSLKSEGERNHIIQLSEVDKTEELTEHILKQL 360
361 TQASLOSINTHENENOKLOKLUKUMTELYQENEMKLRKLTVEENYALEKEEKLKVD 420
Db 361 TQASLOSINTHENENOKLOKLUKUMTELYQENEMKLRKLTVEENYALEKEEKLKVD 420
421 EKISHATEELETYRKAKOLEEELERTIHSYQOQIISHEKKAHDNWLAAARNALNDLR 480
Db 421 EKISHATEELETYRKAKOLEEELERTIHSYQOQIISHEKKAHDNWLAAARNALNDLR 480
481 KENAHNRQKLTETELKELLEKDPYALDVNTAFGREHSPYSGPLGWSSETRAFLSP 540
Db 481 KENAHNRQKLTETELKELLEKDPYALDVNTAFGREHSPYSGPLGWSSETRAFLSP 540
541 TLLEGLRLSLPLPGGGGRSGRPGNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSP 600
Db 541 TLLEGLRLSLPLPGGGGRSGRPGNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSP 600
601 WQDRRMFPPPGQSYDPSALPORQDRFCNSGRSLSGPAELRSFNMPSLDKMDGSMSE 660
Db 601 WQDRRMFPPPGQSYDPSALPORQDRFCNSGRSLSGPAELRSFNMPSLDKMDGSMSE 660
661 MESSRNDTKDDLGNLNVDSLSLPAENATGPGFVPPPLAIPRGPLFPVDARGPFLRRGPP 720
Db 661 MESSRNDTKDDLGNLNVDSLSLPAENATGPGFVPPPLAIPRGPLFPVDARGPFLRRGPP 720
721 FPPPPGAMFGASRDYFPFRDFFGPPAPFAMRNVPYPRGFPYLPFRGFPFPPHSEG 780
Db 721 FPPPPGAMFGASRDYFPFRDFFGPPAPFAMRNVPYPRGFPYLPFRGFPFPPHSEG 780
781 RSEFPGSLIPPSNEPATEHEPQOET 806

Db 781 RGEFSPGLPPSKEPATGHPEQDIT 806

RESULT 13

PCT-US01-08631-35639

; Sequence 35639, Application PC/TUS0108631

; GENERAL INFORMATION:

; APPLICANT: Hysec, Inc

; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

; FILE REFERENCE: 21272-049

; CURRENT APPLICATION NUMBER: PCT/US01/08631

; CURRENT FILING DATE: 2001-03-30

; PRIOR FILING DATE: 2000-03-31

; PRIOR APPLICATION NUMBER: 09/540,217

; PRIOR FILING DATE: 2000-03-31

; PRIOR APPLICATION NUMBER: 09/649,167

; NUMBER OF SEQ ID NOS: 60736

; SOFTWARE: Custom

; SEQ ID NO 35639

; LENGTH: 881

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)...(881)

; OTHER INFORMATION: Xaa = X or * as defined in Table 2

PCT-US01-08631-35639

Query Match 88.4%; Score 3740; DB 1; Length 881;

Best Local Similarity 90.0%; Pred. No. 4.1e-204;

Matches 730; Conservative 22; Mismatches 51; Indels 8; Gaps 7;

Qy 1 MEEPGATPQYLGVLVEELRRVVAALPESMRPDENPYGPPSELVCAAVIGFVVLFLW 60

Db 59 MEEPGATPQYLGVLVEELRRVVAALPEGMRPDENPYGPPSELVCAAVIGFVVLFLW 118

Qy 61 RSFRSVRSRLVYVREOKLGATLSGLIEBKCKLLEKFSLIQKEYEGYEVESSELDASFEKE 120

Db 119 RSFRSVRSRLVYVREOKLGATLSGLIEBKCKLLEKFSLIQKEYEGYEVESSELDASFEKE 178

Qy 121 AAEEARSLEATCEKLNRSNLEDEILCLEKDLKQEKSKHQDELMDADISKIQSLEDE 180

Db 179 AT-EAQSLEATCEKLNRSNLEDEILCLEKDLKQEKSKHQDELMDADISKIQSLEDE 237

Qy 181 SKSLKSOIAEAKICTKTFKMSERRATAIKDALNENSQLOTSKQLFOQEAQVWKGEVSE 240

Db 238 SKSLKSOIAEAKICTKTFKMSERRATAIKDALNENSQLOTSKQLFOQEAQVWKGEVSE 296

Qy 241 LNKOKITTFEDSKVHAQVLDNDKENHIKTLTGHLPMKQDQAAVLEEDTTDDNLELVNSE 300

Db 297 LNKOKITTFEDSKVHAQVLDNDKENHIKTLTGHLPMKQDQAAVLEEDTTDDNLELVNSE 356

Qy 301 SENGAYLDNPPKQALKKLIHAAKLNASLKTLEGERNOIYIQLSEVDKTEELTEHIKNLQ 360

Db 357 SENGAYLDNPPKQALKKLIHAAKLNASLKTLEGERNOIYIQLSEVDKTEELTEHIKNLQ 416

Qy 361 TQASLOSENTHFENENQKLOKLVMTLEYQENEMKLRKLTVEENYRLEKEKLSKVD 420

Db 417 TQASLOSENTHFENENQKLOKLVMTLEYQENEMKLRKLTVEENYRLEKEKLSKVD 476

Qy 421 EKISHATEELETYKRAKDLSEELERTHSYQGOIISHEKKAHDNWLAAARNNDLNR 480

Db 477 EKISHATEELETYKRAKDLSEELERTHSYQGOIISHEKKAHDNWLAAARNNDLNR 536

Qy 481 KENAHNRKQLTETELKFELEKDPYALDVNPTAFGREHSPYGPSPGLWPSSETRAFLSP 540

Db 537 KENAHNRKQLTETELKFELEKDPYALDVNPTAFGREHSPYGPSPGLWPSSETRAFLSP 596

Qy 541 TLLEGPLLSPLPGGGGRGSGPNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSP 600

Db 597 TLLEGPLLSPLPGGGGRGSGPNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSP 656

Qy 601 WQDQRRMMFPQSQSYPSDSALPPQDRFCNSGRISGPAELRSFNNPSLDKMDGSMPSR 660

Db 657 WQDQRRMMFPQSQSYPSDSALPPQDRFCNSGRISGPAELRSFNNPSLDKMDGSMPSR 716

Qy 661 MESSRNDTKDGLNLNVDPSSSLPAENEATGPGFVPPPLAPIRGELFPVDARGPFLRRGPP 720

Db 717 MESSRNDTKDGLNLNVDPSSSLPAENEATGPGFVPPPLAPIRGELFPVDARGPFLRRGPP 776

Qy 721 FPPPPGAMFGASRDYFPPRDF-PGPPAPFAVRN-VYPRGPPPYL--PPRP-GTFPPPP 775

Db 777 FPPPPGAMFGASRDYFPPRDF-PGPPAPFAVRN-VYPRGPPPYL--PPRP-GTFPPPP 836

Qy 776 PH-SEGRSERPSGLIPPSNEPATEHPEPQOE 805

Db 837 PHFXRXEVNFPXVLIPPSNEPATEHPEPQOE 867

RESULT 14

PCT-US01-08631-50617

; Sequence 50617, Application PC/TUS0108631

; GENERAL INFORMATION:

; APPLICANT: Hysec, Inc

; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

; FILE REFERENCE: 21272-049

; CURRENT APPLICATION NUMBER: PCT/US01/08631

; CURRENT FILING DATE: 2001-03-30

; PRIOR FILING DATE: 2000-03-31

; PRIOR APPLICATION NUMBER: 09/540,217

; PRIOR FILING DATE: 2000-03-31

; PRIOR APPLICATION NUMBER: 09/649,167

; NUMBER OF SEQ ID NOS: 60736

; SOFTWARE: Custom

; SEQ ID NO 50617

; LENGTH: 881

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: DOMAIN

; LOCATION: (514)...(568)

; OTHER INFORMATION: eMATRIX, accession number BL011608, p-value=8.551e-09, raw score (

; OTHER INFORMATION: 19.54

; NAME/KEY: misc feature

; LOCATION: (1)...(881)

; OTHER INFORMATION: Xaa = X or * as defined in Table 2

PCT-US01-08631-50617

Query Match 87.7%; Score 3712; DB 1; Length 881;

Best Local Similarity 89.4%; Pred. No. 1.6e-202;

Matches 725; Conservative 24; Mismatches 54; Indels 8; Gaps 7;

Qy 1 MEEPGATPQYLGVLVEELRRVVAALPESMRPDENPYGPPSELVCAAVIGFVVLFLW 60

Db 59 MEEPGATPQYLGVLVEELRRVVAALPEGMRPDENPYGPPSELVCAAVIGFVVLFLW 118

Qy 61 RSFRSVRSRLVYVREOKLGATLSGLIEBKCKLLEKFSLIQKEYEGYEVESSELDASFEKE 120

Db 119 RSFRSVRSRLVYVREOKLGATLSGLIEBKCKLLEKFSLIQKEYEGYEVESSELDASFEKE 178

Qy 121 AAEEARSLEATCEKLNRSNLEDEILCLEKDLKQEKSKHQDELMDADISKIQSLEDE 180

Db 179 AT-EAQSLEATCEKLNRSNLEDEILCLEKDLKQEKSKHQDELMDADISKIQSLEDE 237

Qy 181 SKSLKSOIAEAKICTKTFKMSERRATAIKDALNENSQLOTSKQLFOQEAQVWKGEVSE 240

Db 238 SKSLKSOIAEAKICTKTFKMSERRATAIKDALNENSQLOTSKQLFOQEAQVWKGEVSE 296

Qy 241 LNKOKITTFEDSKVHAQVLDNDKENHIKTLTGHLPMKQDQAAVLEEDTTDDNLELVNSE 300

Db 297 LNKOKITTFEDSKVHAQVLDNDKENHIKTLTGHLPMKQDQAAVLEEDTTDDNLELVNSE 356

Qy 301 SENGAYLDNPPKQALKKLIHAAKLNASLKTLEGERNOIYIQLSEVDKTEELTEHIKNLQ 360

Db 357 SENGAYLDNPPKQALKKLIHAAKLNASLKTLEGERNOIYIQLSEVDKTEELTEHIKNLQ 416

Qy 361 TQASLOSENTHFENENQKLOKLVMTLEYQENEMKLRKLTVEENYRLEKEKLSKVD 420

Db 417 TQASLOSENTHFENENQKLOKLVMTLEYQENEMKLRKLTVEENYRLEKEKLSKVD 476

Qy 421 EKISHATEELETYKRAKDLSEELERTHSYQGOIISHEKKAHDNWLAAARNNDLNR 480

Db 477 EKISHATEELETYKRAKDLSEELERTHSYQGOIISHEKKAHDNWLAAARNNDLNR 536

Qy 481 KENAHNRKQLTETELKFELEKDPYALDVNPTAFGREHSPYGPSPGLWPSSETRAFLSP 540

Db 537 KENAHNRKQLTETELKFELEKDPYALDVNPTAFGREHSPYGPSPGLWPSSETRAFLSP 596

Qy 541 TLLEGPLLSPLPGGGGRGSGPNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSP 600

Db 597 TLLEGPLLSPLPGGGGRGSGPNPLDHOITNERGESSCDRLTDPHRAPSDTGSLSP 656

Db 357 SENGAYLDNPPKALKKLIHAAKLNASLKTLEGERNQIYIQLSEVDKTEELTEHKNLQ 416
Qy 361 TQASLQSENTHFENENKLOKLVMTLEYQENEMKLRKLTVEENYLEKEEKLKVD 420
Db 417 TEQASLQSENTHFENENKLOKLVMTLEYQENEMKLRKLTVEENYLEKEEKLKVD 476
Qy 421 EKISHATEELETYRKAKOLEBEELERTIHSYQOQIISHEKKAHDNWLAAARNAERLNDLR 480
Db 477 EKISHATEELETYRKAKOLEBEELERTIHSYQOQIISHEKKAHDNWLAAARNAERLNDLR 536
Qy 481 KENAHNRQKLTETELKPELLEKDPYALDVPNTAFGREHSPYGPSPGLGWSSSTRAFLSPP 540
Db 537 KENAHNRQKLTETELKPELLEKDPYALDVPNTAFGREHSPYGPSPGLGWSSSTRAFLSPP 596
Qy 541 TLLEGPLRLSPLLPGGGGRSGRGNPLDQIITNERGESSCDRLTDPHRAPSDGTGSLSP 600
Db 597 TLLEGPLRLSPLLPGGGGRSGRGNPLDQIITNERGESSCDRLTDPHRAPSDGTGSLSP 656
Qy 601 WDQDRMMFPFPGQSPYDPSALPQQRDFCSNGLSGPAELRSFNMPSLDKMDGSMPS 660
Db 657 WDQDRMMFPFPGQSPYDPSALPQQRDFCSNGLSGPAELRSFNMPSLDKMDGSMPS 716
Qy 661 MESSRNDTKDDLGNLVNPDSSSLPAENEATGFGVPPPLAPIRGPLEFVDPARGPFLRRGPP 720
Db 717 MESSRNDTKDDLGNLVNPDSSSLPAENEATGFGVPPPLAPIRGPLEFVDPARGPFLRRGPP 776
Qy 721 FPPPPGAMFGASRDYPPPPDF-FGPPPPAPFAMRN-VYPPRGPPPYL--PPRP-GFPFPP 775
Db 777 FPPPPGAMFGASRDYPPPPDF-FGPPPPAPFAMRN-VYPPRGPPPYL--PPRP-GFPFPP 836
Qy 776 PH-SEGSRSEPPGLIPPSNEPATEHPEPQOE 805
Db 837 PHFXRXEVNPPXVLIPTSNEPATEHPEPQOE 867

RESULT 15
US-10-236-177-180
; Sequence 180, Application US/10236177
; GENERAL INFORMATION:
; APPLICANT: Alsbrook, John P
; APPLICANT: Anderson, David W
; APPLICANT: Bergths, Constance
; APPLICANT: Burgess, Catherine E
; APPLICANT: Catterton, Elina
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Ji, Weizhen
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Leach, Martin D
; APPLICANT: Lepley, Denise M
; APPLICANT: Li, Li
; APPLICANT: Liu, Xiaohong
; APPLICANT: Miller, Charles E
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Reiger, Daniel K
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Shenoy, Suresh
; APPLICANT: Shimkets, Richard A
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Taupier, Raymond J
; APPLICANT: Vernier, Corine AM
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 21402-442D
; CURRENT APPLICATION NUMBER: US/10/236,177
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 60/318,120
; PRIOR FILING DATE: 2001-09-07

; PRIOR APPLICATION NUMBER: 60/318,130
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/318,219
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/318,430
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/322,781
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/322,636
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/322,817
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/373,212
; PRIOR FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 60/322,816
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/323,519
; PRIOR FILING DATE: 2001-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 434
; SOFTWARE: Curasequest version 0.1
; SEQ ID NO 180
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-236-177-180
Query Match 87.6%; Score 3705; DB 28; Length 760;
Best Local Similarity 88.7%; Pred. No. 3.3e-202;
Matches 715; Conservative 19; Mismatches 26; Indels 46; Gaps 3;
Qy 1 MEEPGATPQPYLGVLBELRRVVAALPESMRPDENPYGFPSELVCAAVIGFFVVLFLW 60
Db 1 MEEPGATPQPYLGVLBELRR----- 21
Qy 61 RSFRSVRSRLVYGRQKLGATLSGLIEKCKLLEKFSLIQKEYGYEVESSELEDAFPEKE 120
Db 22 -----VRSRLVYGRKELALMLSGLIEKCKLLEKFSLIQKEYGYEVESSELEDAFPEKE 76
Qy 121 AAEEARSLEATCEKLNRSNLEDEILCLEKDLKCKSKSQOQELMADISKSIOSEDE 180
Db 77 AT-EAQSLEATCEKLNRSNLEDEILCLEKDLKCKSKSQOQELMADISKSIOSEDE 135
Qy 181 SKSLKSQIAEAKIICKTFKMSERRAIAIKDALNENSQLOTSKQLFOEAFAVWKEVSE 240
Db 136 SKSLKSQIAEAKMTFKI FOMNEERLKI AIKDALNENSQLOESQKOLL-QEAEVWKEVSE 194
Qy 241 LNKQKITPDSKVHAEQVLDNKENHIKTLTGHLPMKQDAVLBEDTTDDNLELVNSE 300
Db 195 LNKQKITPDSKVHAEQVLDNKENHIKTLTGHLPMKQDAVLBEDTTDDNLELVNSE 254
Qy 301 SENGAYLDNPPKALKKLIHAAKLNASLKTLEGERNQIYIQLSEVDKTEELTEHKNLQ 360
Db 255 SENGAYLDNPPKALKKLIHAAKLNASLKTLEGERNQIYIQLSEVDKTEELTEHKNLQ 314
Qy 361 TQASLQSENTHFENENKLOKLVMTLEYQENEMKLRKLTVEENYLEKEEKLKVD 420
Db 315 TEQASLQSENTHFENENKLOKLVMTLEYQENEMKLRKLTVEENYLEKEEKLKVD 374
Qy 421 EKISHATEELETYRKAKOLEBEELERTIHSYQOQIISHEKKAHDNWLAAARNAERLNDLR 480
Db 375 EKISHATEELETYRKAKOLEBEELERTIHSYQOQIISHEKKAHDNWLAAARNAERLNDLR 434
Qy 481 KENAHNRQKLTETELKPELLEKDPYALDVPNTAFGREHSPYGPSPGLGWSSSTRAFLSPP 540
Db 435 KENAHNRQKLTETELKPELLEKDPYALDVPNTAFGREHSPYGPSPGLGWSSSTRAFLSPP 494
Qy 541 TLLEGPLRLSPLLPGGGGRSGRGNPLDQIITNERGESSCDRLTDPHRAPSDGTGSLSP 600
Db 495 TLLEGPLRLSPLLPGGGGRSGRGNPLDQIITNERGESSCDRLTDPHRAPSDGTGSLSP 554
Qy 601 WDQDRMMFPFPGQSPYDPSALPQQRDFCSNGLSGPAELRSFNMPSLDKMDGSMPS 660

Db 555 WQDRRMFPFGOSYPDSALPPQDRFCNSGRLSGPAELRSFNNMPSLDKMDGMPSE 614
QY 661 MESSRNDTKDDLGNLNVDPDSSLPAENEATGPGFVPPPLAPIRGGLFPVDARGPFLRRGPP 720
Db 615 MESSRNDTKDDLGNLNVDPDSSLPAENEATGPGFVPPPLAPVRGGLFPVDARGPFLRRGPP 674
QY 721 FPPPPPGAMFGASRDYFPPRRDPGPPAPAFAMRNVPYPPRGFPYLPDRPGFPPPPHSEG 780
Db 675 FPPPPPGAMFGASRDYFPPGGDFGPPAPAFAMRNVPYPPRGFPYLPDRPGFPPPPHSEG 734
QY 781 RSEPPSGLIPPSNEPATEHPEPQOET 806
Db 735 RSEPPSGLIPPSNEPATEHPEPQOET 760

Search completed: April 20, 2004, 07:58:47
Job time : 191 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2004, 07:52:48 ; Search time 18 Seconds
(without alignments);
1235.532 Million cell updates/sec

Title: US-09-744-314-6

Perfect score: 4231

Sequence: 1 MEFGATPQYLGVLVLELR.....GLIPPSNEPATEHPPOQET 806

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 196679 seqs, 27592530 residues

Total number of hits satisfying chosen parameters: 196679

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA New:*

- 1: /cgn2_6/prodata/1/paa/pct_NEW COMB.pep.*
- 2: /cgn2_6/prodata/1/paa/US05_NEW COMB.pep.*
- 3: /cgn2_6/prodata/1/paa/US07_NEW COMB.pep.*
- 4: /cgn2_6/prodata/1/paa/US08_NEW COMB.pep.*
- 5: /cgn2_6/prodata/1/paa/US09_NEW COMB.pep.*
- 6: /cgn2_6/prodata/1/paa/US10_NEW COMB.pep.*
- 7: /cgn2_6/prodata/1/paa/US60_NEW COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3946.5	93.3	763	US-10-486-020-24	Sequence 24, Appl
2	3900	92.2	804	US-10-486-020-11	Sequence 11, Appl
3	270	6.4	3114	PCT-US04-00035-99	Sequence 99, Appl
4	256	6.1	1684	PCT-US04-07412-885	Sequence 885, Appl
5	250.5	5.9	496	US-10-767-701-43815	Sequence 43815, A
6	250	5.9	2325	US-10-663-433-2	Sequence 2, Appl
7	249.5	5.9	1960	US-60-555-303-377	Sequence 377, Appl
8	249.5	5.9	1960	US-60-555-303-378	Sequence 378, Appl
9	249.5	5.9	1960	US-60-555-303-379	Sequence 379, Appl
10	249.5	5.9	1960	US-60-555-303-380	Sequence 380, Appl
11	247.5	5.8	1958	US-10-791-666-4	Sequence 4, Appl
12	247.5	5.8	2054	US-10-791-666-2	Sequence 2, Appl
13	245	5.8	588	PCT-US04-02188-109	Sequence 109, Appl
14	245	5.8	588	US-10-764-425-109	Sequence 109, Appl
15	245	5.8	1682	US-10-796-307-578	Sequence 578, Appl
16	245	5.8	1938	US-10-796-307-576	Sequence 576, Appl
17	245	5.8	1938	US-10-796-307-577	Sequence 577, Appl
18	245	5.8	1954	US-10-796-307-575	Sequence 575, Appl
19	245	5.8	1972	US-10-796-307-579	Sequence 579, Appl
20	245	5.8	1972	US-10-796-307-581	Sequence 581, Appl
21	245	5.8	2871	US-60-552-390-194	Sequence 194, Appl
22	245	5.8	2871	US-60-552-390-195	Sequence 195, Appl
23	243.5	5.8	1235	US-10-784-004-783	Sequence 783, Appl
24	243.5	5.8	1235	US-10-784-004-1111	Sequence 1111, Appl
25	243.5	5.8	1305	US-10-784-004-671	Sequence 671, Appl
26	243.5	5.8	1305	US-10-784-004-789	Sequence 789, Appl

Sequence 1066, Ap
Sequence 1114, Ap
Sequence 754, App
Sequence 1098, Ap
Sequence 699, App
Sequence 1078, Ap
Sequence 668, App
Sequence 1064, Ap
Sequence 659, App
Sequence 1058, Ap
Sequence 662, App
Sequence 1061, Ap
Sequence 633, App
Sequence 650, App
Sequence 652, App
Sequence 757, App
Sequence 781, App
Sequence 1040, Ap
Sequence 1053, Ap

ALIGNMENTS

RESULT 1

US-10-486-020-24
; Sequence 24 Application US/10486020
; GENERAL INFORMATION: Valda; BARROSO, Ines;
; APPLICANT: AZIMZAL, Mariah R.; BECHA, Shanya D.;
; APPLICANT: BOROWSKY, Mark L.; DUGGAN, Brendan M.;
; APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.;
; APPLICANT: FORSYTHE, Ian J.; GIETZEN, Kimberly J.;
; APPLICANT: GORVAD, Ann B.; GRAUL, Richard C.;
; APPLICANT: GRIFFIN, Jennifer A.; GURURAJAN, Rajagopal;
; APPLICANT: HAPALIA, April J.A.; ISON, Craig H.;
; APPLICANT: KABLE, Amy E.; KHAN, Farrah A.;
; APPLICANT: LEE, Sally; LEE, Soo Yeun;
; APPLICANT: LI, Joana X.; REDDY, Roopa;
; APPLICANT: RICHARDSON, Thomas W.; SPRAGUE, William W.;
; APPLICANT: SWARNAKAR, Anita; TANG, Y. Tom;
; APPLICANT: WARREN, Bridget A.; XU, Yuming;
; APPLICANT: YAO, Monique G.; YUE, Henry;
; APPLICANT: YUE, Huibin
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PF-1126 USN
; CURRENT APPLICATION NUMBER: US/10/486,020
; PRIOR FILING DATE: 2004-02-05
; PRIOR APPLICATION NUMBER: PCT/US02/25465
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/311,017
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/313,070
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/313,071
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/314,678
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/316,692
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/317,913
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/322,182
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/340,747
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/342,761
; PRIOR FILING DATE: 2001-12-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PERL Program
; SEQ ID NO 24
; LENGTH: 763

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7506167CD1
US-10-486-020-24

Query Match 93.3%; Score 3946.5; DB 6; Length 763;
Best Local Similarity 94.3%; Pred. No. 5.9e-156;
Matches 760; Conservative 1; Mismatches 2; Indels 43; Gaps 1;

QY 1 MEEPGATPQYLGVLVLEELRRVVAALPESMRPDENPYGPPSELVCAAVIGFFVVLFLW 60
DB 1 MEEPGATPQYLGVLVLEELRRVVAALPESMRPDENPYGPPSELVCAAVIGFFVVLFLW 60

QY 61 RSFRSVRSRLYVGRQKLGATLSGLIEBKCKLLEKFSLIQKEYEGYEVESSELDASFEKE 120
DB 61 RSFRSVRSRLYVGRQKLGATLSGLIEBKCKLLEKFSLIQKEYEGYEVESSELDASFEKE 120

QY 121 AAEARSLEATCEKLNRSNSELDEILCLEKDKQKSKHSQOQDELMDISKSIQSLDE 180
DB 121 AAEARSLEATCEKLNRSNSELDEILCLEKDKQKSKHSQOQDELMDISKSIQSLDE 180

QY 181 SKSLKSQIAEAKIICTKTFKMSERRAIAIKDALNENSQLOTSHKQFQQAENVKGEVSE 240
DB 181 SKSLKSQIAEAKIICTKTFKMSERRAIAIKDALNENSQLOTSHKQFQQAENVKGEVSE 240

QY 241 LNKOKITFEDSKVHAEOVLNDKENHIKTLTGHLPMKQQAALVEEDTTDDNLELVNSE 300
DB 241 LNKOKITFEDSKVHAEOVLNDKENHIKTLTGHLPMKQQAALVEEDTTDDNLELVNSE 300

QY 301 SENGAYLDNPPKGAUKKIIHAALKNASLKTLEGERNQIYIQLSEVDKTEKEITHIKNLQ 360
DB 301 WENGAYLDNPPKGAUKKIIHAALKNASLKTLEGERNQIYIQLSEVDKTEKEITHIKNLQ 360

QY 361 TQASLOSNTFHNENKQLOKQKLVMTELYOENEMKLRKLTVEENYRLKEEKLKVD 420
DB 361 TQASLOSNTFHNENKQLOKQKLVMTELYOENEMKLRKLTVEENYRLKEEKLKVD 420

QY 421 EKISHATELEYRKAADLEELERTHSYOGQIISHEKKAHDNWLAAARNAERNLNDLR 480
DB 421 EKISHATELEYRKAADLEELERTHSYOGQIISHEKKAHDNWLAAARNAERNLNDLR 480

QY 481 KENAHNRQKLTETELKFELEKDPYALDVPNTAFGREHSFYGPSPLGWPSPSETRAFSLPP 540
DB 481 KENAHNRQKLTETELKFELEKDPYALDVPNTAF----- 514

QY 541 TLLEGSLRLSPLLPGGGGRSGRGNPLDHOITNERGESSCDRLTDHRAPSDTCGLSPP 600
DB 515 -----GRGSRGPNPLDHOITNERGESSCDRLTDHRAPSDTCGLSPP 557

QY 601 WQDRRMVFPFPGQSYDPSALPPQDRFCNSGRSLGPAELRSFNMPSLDKWDGSMFSE 660
DB 558 WQDRRMVFPFPGQSYDPSALPPQDRFCNSGRSLGPAELRSFNMPSLDKWDGSMFSE 617

QY 661 MESSRNDTKDDLGNLVDPSSLPAAENATGPGFVPPPLAPIRGBLFFVDARGFELRRGPP 720
DB 618 MESSRNDTKDDLGNLVDPSSLPAAENATGPGFVPPPLAPIRGBLFFVDARGFELRRGPP 677

QY 721 FPPPPFGAGASRDYFPPDRFPFPPAPFAMRNVPYPRGFPYLPFRFPFPPPHSEG 780
DB 678 FPPPPFGAGASRDYFPPDRFPFPPAPFAMRNVPYPRGFPYLPFRFPFPPPHSEG 737

QY 781 RSEFPGLIPPSNEPATEPEPOQET 806
DB 738 RSEFPGLIPPSNEPATEPEPOQET 763

RESULT 2
US-10-486-020-11
; Sequence 11, Application US/10486020
; GENERAL INFORMATION:
; APPLICANT: AZIMZAI, Yalda; BARROSO, Ines;
; APPLICANT: BAUGHN, Mariah R.; BECHA, Shanya D.;
; APPLICANT: BOROWSKY, Mark L.; DUGGAN, Brendan M.;
; APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.;
; APPLICANT: FORSYTHE, Ian J.; GIETZEN, Kimberly J.;
; APPLICANT: GORVAD, Ann E.; GRAUL, Richard C.;
; APPLICANT: GRPEIN, Jennifer A.; GURURAJAN, Rajagopal;
; APPLICANT: HAFALIA, April J.A.; ISON, Craig H.;
; APPLICANT: KABLE, Amy E.; KHAN, Farrah A.;
; APPLICANT: LEE, Sally; LEE, Soo Yeun;
; APPLICANT: LI, Joana X.; REDDY, Roopa;
; APPLICANT: RICHARDSON, Thomas W.; SPRAGUE, William W.;
; APPLICANT: SWARNAKAR, Anita; TANG, Y. Tom;
; APPLICANT: WARREN, Bridget A.; XU, Yuming;
; APPLICANT: YAO, Monique G.; YUE, Henry;
; APPLICANT: YUE, Huidun
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PF-1126 USN
; CURRENT APPLICATION NUMBER: US/10/486,020
; CURRENT FILING DATE: 2004-02-05
; PRIOR APPLICATION NUMBER: PCT/US02/25465
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/311,017
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/313,070
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/313,071
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/314,678
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/316,692
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/317,913
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/322,182
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/340,747
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/342,761
; PRIOR FILING DATE: 2001-12-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PERL Program
; SEQ ID NO 11
; LENGTH: 804
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 6715627CD1
US-10-486-020-11

Query Match 92.8%; Score 3900; DB 6; Length 804;
Best Local Similarity 92.8%; Pred. No. 5.5e-156;
Matches 748; Conservative 20; Mismatches 36; Indels 2; Gaps 2;

QY 1 MEEPGATPQYLGVLVLEELRRVVAALPESMRPDENPYGPPSELVCAAVIGFFVVLFLW 60
DB 1 MEEPGATPQYLGVLVLEELRRVVAALPESMRPDENPYGPPSELVCAAVIGFFVVLFLW 60

QY 61 RSFRSVRSRLYVGRQKLGATLSGLIEBKCKLLEKFSLIQKEYEGYEVESSELDASFEKE 120
DB 61 RSFRSVRSRLYVGRQKLGATLSGLIEBKCKLLEKFSLIQKEYEGYEVESSELDASFEKE 120

QY 121 AAEARSLEATCEKLNRSNSELDEILCLEKDKQKSKHSQOQDELMDISKSIQSLDE 180
DB 121 AT-EAQSLEATCEKLNRSNSELDEILCLEKDKQKSKHSQOQDELMDISKSIQSLDE 179

QY 181 SKSLKSQIAEAKIICTKTFKMSERRAIAIKDALNENSQLOTSHKQFQQAENVKGEVSE 240
DB 180 SKSLKSQIAEAKIICTKTFKMSERRAIAIKDALNENSQLOTSHKQFQQAENVKGEVSE 238

QY 241 LNKOKITFEDSKVHAEOVLNDKENHIKTLTGHLPMKQQAALVEEDTTDDNLELVNSE 300
```

```
Db 239 LNKQKVTBDSKVHAEQVINDKESHKILTERLLKXKDWAAAMLGEDITDDDDNLEMMNSE 298
Qy 301 SENGAYLDNPPGALKLTHAAKLNASKLTGBERNQYIQLSEVDKTEBELTEHKNLQ 360
Db 299 SENGAYLDNPPGALKLTHAAKLNASKLTGBERNQYIQLSEVDKTEBELTEHKNLQ 358
Qy 361 TQOASLQSENTHFENENQKLOKLVMTLYQENEMKLRKLTVEENYRLEKEKLSKVD 420
Db 359 TQOASLQSENTHFENENQKLOKLVMTLYQENEMKLRKLTVEENYRLEKEKLSKVD 418
Qy 421 EKISHATELETRYKRAKDLSEBELTTHSYOGQIISHKKAHDNWLAAARNAERLNDLR 480
Db 419 EKISHATELETRYKRAKDLSEBELTTHSYOGQIISHKKAHDNWLAAARNAERLNDLR 478
Qy 481 KENAHNRQKLTETELKFELEKDPYALDVPNTAFGRHSYPGSPGLWPMSPSTRFLSP 540
Db 479 KENAHNRQKLTETELKFELEKDPYALDVPNTAFGRHSYPGSPGLWPMSPSTRFLSP 538
Qy 541 TLLEGRLRLSPLLPGGGGSGRPGNPLDQITNERGSSCDRLTDPHPAPSDTGLSP 600
Db 539 TLLEGRLRLSPLLPGGGGSGRPGNPLDQITNERGSSCDRLTDPHPAPSDTGLSP 598
Qy 601 WDQDRMMPPPQSYDPSALPQDRFCNSGRLSGPAELRSFNMPSLDXDGMSPSE 660
Db 599 WDQDRMMPPPQSYDPSALPQDRFCNSGRLSGPAELRSFNMPSLDXDGMSPSE 658
Qy 661 MESSRNDTKDGLNLPDSSLPAPENATGPGFVPPPLAPIGPLPPVDARGPFLRGPP 720
Db 659 MESSRNDTKDGLNLPDSSLPAPENATGPGFVPPPLAPIGPLPPVDARGPFLRGPP 718
Qy 721 FPPPPGAMFGASRDYFPDRDPGPPAPFAMNVYPPRGFPYLPFRGFPFPPHSG 780
Db 719 FPPPPGAMFGASRDYFPDRDPGPPAPFAMNVYPPRGFPYLPFRGFPFPPHSG 778
Qy 781 RSEFPGLIPPSNEPATEHPEPOQET 806
Db 779 RSEFPGLIPPSNEPATEHPEPOQET 804

RESULT 3
PCT-US04-00035-99
; Sequence 99, Application PC/TUS0400035
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Martinez, Robert
; APPLICANT: Brown, Eugene
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING AND TREATING COLON
; FILE REFERENCE: AM100927 (031836-002000)
; CURRENT APPLICATION NUMBER: PCT/US04/00035
; PRIOR FILING DATE: 2004-01-06
; PRIOR APPLICATION NUMBER: US Provisional Application 60/438,000
; NUMBER OF SEQ ID NOS: 54873
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 99
; LENGTH: 3114
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US04-00035-99

Query Match 6.4%; Score 270; DB 1; Length 3114;
Best Local Similarity 21.1%; Pred. No. 0.00062;
Matches 124; Conservative 110; Mismatches 185; Indels 170; Gaps 22;

Qy 76 QKLGATSLGIEEKKLLEKFLSIQKEVEGY-----EVESSLEDASPEK-AAEARSLE 129
Db 2017 QTLSSDVSLLKDKTHLQKLSLEKDSQALSUTKCELEQNIAQLNKEKLLVKESESUQ 2076
Qy 130 ATC-----EKLNRN-----SELEIDLCLEKDLKOKSKHSQQDELWA 168
Db 2077 ARLSSEYKLVNKALEAALVEKGEFALRLSS*QEEVHQLRGIEKLVRIRIEADEKQOL 2136
```

```
Qy 169 DISKSIQSLDESKLSQIAEAKIICKTFKMSERRAIRAKDALNENSQLOTF----- 221
Db 2137 HIAEKLKERENDSLKDKVENLE---RELQMSSENOELVILDAENSKAEVETLKTQIEE 2193
Qy 222 -----SHKOLFQOEAENVKGEVSELNKKQKITF-----ED 250
Db 2194 MARSLKVFEIDLVTLRSEKENLTQIQOEGQQLSELDKLLSSFKSLLEKEQAEIQIEE 2253
Qy 251 SKVHAEQVLN-----DKE-----NHIKTLTGLHPLM- 275
Db 2254 SKTAVENLQNLKELNEAVALCGDQELMKATQSLDPPTEEBEHQLRNSIEKLRLARLEAD 2313
Qy 276 MKQAAVLBEDTDDD-----DNLELEV---NSESENGAYLDNPPGALKLTHAA 322
Db 2314 EKQOLCVLOQLKESSEHADLLKGRVENLERELRIARTNQEHAALEAENSKEVETL--KA 2371
Qy 323 KINASLTGLGERNQIYIQLSEVDKTEBELTEHKNLQOQASLQSENTHFEN-----EN 377
Db 2372 KIEGWTQSLRG-----LELDVVVTIRSEKENLTNLQKEQERISELEINSSFENILQKEQ 2427
Qy 378 QKLOQKLVMT--ELYOENEMKLRKLTV-----EEN-----YRLEKEKLS 417
Db 2428 EKQMKESKSTAMEMLQTLKELNERNVAALHNDQEAACKAEQNLSSQVECLEKEQAQLLQ 2487
Qy 418 KVDB-----KISHATELETRYKRAKDLSEBELTTHSYOGQ-----IISHEKKA 462
Db 2488 GLDEAKNYYIVLQSSVNGLIQEVEDGKQLEKKDEEISRLKNQIQDQEQVLVSKLSQVEGE 2547
Qy 463 HDNW-----LAARN-----ABRNLDLRKENAHNRKLTETELKFELEK 503
Db 2548 HQLWKEQNLELRNLTVLEQKIQVLQSKNASLQDTLEVLQSSYKNLENE 2596

RESULT 4
PCT-US04-07412-885
; Sequence 885, Application PC/TUS0407412
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Wang, Jian-rui
; APPLICANT: Zhang Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Ghosh, Malabika
; APPLICANT: Xue, Aidong J.
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Dunrui
; APPLICANT: Goodrich, Kyle W.
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Gezhi
; APPLICANT: Wang, Zhiwei
; APPLICANT: Boyle, Bryan J.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 822CIP/PCT
; CURRENT APPLICATION NUMBER: PCT/US04/07412
; CURRENT FILING DATE: 2004-03-19
; PRIOR APPLICATION NUMBER: US 10/389,559
; PRIOR FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: US 60/365,264
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: US 60/340,187
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 10/296,115
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: PCT/US00/35017
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/488,725
```

```
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 10/275,027
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1920
; SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 885
; LENGTH: 1684
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US04-07412-885

Query Match
Best Local Similarity 6.1%; Score 256; DB 1; Length 1684;
Matches 115; Conservative 103; Mismatches 187; Indels 106; Gaps 22;

QY 74 REOKL---GATLSGLIECKKLEKFSLIQKEYEGYEVESSELESAS-----FEKAAAEAR 126
Db 1110 KEQKIKHEATTVELEEVQVQCKKQLQKTMQ--EELVKKDAQQTTLNMEIADYER 1167
QY 127 SLEATEKLNRSNLEDEILCKDLQKSKHSDQDELMDADISKSIQSLEDESLSKS 186
Db 1168 LMKELNQLTKNNKIED---LEQEIKIQKQETLOBEITSLOSSVQYVEEKTKIK- 1222
QY 187 QIAEAKIICKTPESEERRAIAIKALNENSQLOTSK-----KQLSQOEAENVWG 236
Db 1223 -----QLLVKT-----KKELADSKQAETDHLILQASLGELEASQQQVEVYKI 1265
QY 237 EYSELNKKQITPDSKVHAEQVINDKENHIKTLTGLHPMKDQAAVLBEDTDDNLE 296
Db 1266 QLAE-----ITSEKIKH-EHLKTSAEQHQRTLSAY---QQRVTALQEECBAAAEQAT 1315
QY 297 VNSESENGAYLONPPKGLKLIH---AAKLNASKLTKEGE---RNQYIQLSEVDKTYKE 350
Db 1316 VTSEPE-----SYKVRVHNLKKNKSKMSQAETEGAKQERHEML-IDQLKI 1363
QY 351 ELTEHIKNIQ---TQASLSQSE-NTHFENENKLOQKLVMTLEYQENEMKLRKL-TVE 405
Db 1364 KLODSQNNLIQINVELSELOTLQSEHDTLLERHKNLQETV-----SKEAELEKLSIQ 1415
QY 406 ENYRLEKEKLSKVDKIKSHATELETYKRAKDLSEELERTIHSYQGOITISHEK----- 460
Db 1416 SENMMKSEHTQTVSQTSSQNEVLNRSFRDQVRLHQLQEEHRTKVTETLQQQLSKVEAQLFQL 1475
QY 461 KAHNWLAAARNARNINDLRKENAHNRQKLTETELKF-----ELLEKDPVALDVPNTAFGR 516
Db 1476 KNEPTTRSPVSSQSLKXNLRER-----NTDPLLDHMTVTRREGCGMETTTOTESVS 1527
QY 517 EHSYPGSPPLGWSPSSSETRAFLSPPTLLEGPL 547
Db 1528 SASY-----TQSLEQLNSPPTKLEPPL 1551

RESULT 5
US-10-767-701-43815
; Sequence 43815, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 39-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 43815
; LENGTH: 496
; TYPE: PRT
```

```
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C3093_1.pep
US-10-767-701-43815

Query Match
Best Local Similarity 5.9%; Score 250.5; DB 6; Length 496;
Matches 124; Conservative 88; Mismatches 134; Indels 139; Gaps 27;

QY 85 LIBECKKLEK-FSLIQKEYEGYEVESSELESASFEKAAAEARSLEATECKLNRSNLE 143
Db 52 LAQNNKEMEKHISLSQEEVKGH-----QDKATDHQIIEE-SLRSTISEL-----KV 98
QY 144 DEILCLEK---DLKOEKSKHQDELMDADISK-----SIQSLDESLSKLSQIA 189
Db 99 QEALSLSKQVADLEQ---KLASQD---ADISKLTLEELNHLHCSSSESLEKTKILETET 152
QY 190 -----EAKIIC---KTFKVSERRAIAIKDAL--NENSQLOTSKQL--FOEAEVWG 236
Db 153 TALEELQAKLLSQEMETKLDQSKGRQTSAALEKQNGQLIVLQAEKLNLDENETLQ 212
QY 237 EYSELNKKQITPDSKVHAEQVINDKENHIKTLTGLHPMKDQAAVLBEDTDDNLE 295
Db 213 SLADLN-SKFSKDSMLHQAEDELAKQ-----LVLSEALSQKELEL 254
QY 296 EVNSES-----ENGAYLONPPKGLK-----KLTHAAK--LN-----ASLKTLEGRNQ 337
Db 255 NLNSLSQEGESKAFGENASQKILELEAQVHAMHAAEALNLELKEAEASVKAENKSSD 314
QY 338 IYIQLSEVDK---TKEELTEHIKNIQTOQASLOSSE-----THFE----- 374
Db 315 LEQQLSEIENKLVASSEEI-ELLKERIQEAAVSAERGMLQEBETMTSVGEYKEKITELQS 373
QY 375 -----NENKLOQKLVMTLEYQENEMKLN-----RKLTVENYRLEK----- 412
Db 374 SLDSSVSKNQLLEQEVKELTDKCEHQEAHSVQQRSELEDLHTSKTHAEGHSRTQE 433
QY 413 -EKLKVDKIKSHATELETYKRAKDLSEELERTIHSYQGOITISHEKKAHNLWLAARN 471
Db 434 LEQELANNYTKLXGVEELEQYERKASQLADDLE-----AYQTKAASLEAVVE-----AASE 485
QY 472 AERNL 476
Db 486 KEKEL 490

RESULT 6
US-10-663-433-2
; Sequence 2, Application US/10663433
; GENERAL INFORMATION:
; APPLICANT: Doreay, Stephen J.
; TITLE OF INVENTION: CENTROSOME PROTEINS AND USES THEREOF
; FILE REFERENCE: 07917-162001
; CURRENT APPLICATION NUMBER: US/10/663,433
; CURRENT FILING DATE: 2003-09-15
; PRIOR APPLICATION NUMBER: US 60/410,520
; PRIOR FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2325
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-663-433-2

Query Match
Best Local Similarity 5.9%; Score 250; DB 6; Length 2325;
Matches 171; Conservative 129; Mismatches 297; Indels 188; Gaps 36;

QY 73 GREQKLGATISG-----LIBECKKLEKFSLIQKEYEGYEVESSELESASFEKAA 121
Db 620 GLQEVLG-TIKGQATQAGNECKLRDEKTLQRLTEVEQERDQLEI-VAMDAENMKEL 677
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2004, 02:44:16 ; Search time 7714 Seconds
(without alignments)
12248.372 Million cell updates/sec

Title: US-09-744-314-14

Perfect score: 3164

Sequence: 1 aggtttaatccatgaagaag.....agttgatttaaaaaaaaaa 3164

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estba.*

2: em_esthum.*

3: em_estinv.*

4: em_estm.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_hci.*

9: gb_est1.*

10: gb_est2.*

11: gb_hci.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_mam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_phg.*

27: em_gss_vri.*

28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2654	83.9	2972	11	BC039017
2	2191.2	69.3	2666	11	BC030655
3	2001.8	63.3	2415	29	AY409371
4	1694	53.5	2415	29	AY409372

RESULT 1
BC039017
LOCUS
DEFINITION
Homo sapiens meningoeloma expressed antigen 6 (coiled-coil
proline-rich), mRNA (cDNA clone IMAGE:5770231), containing
frame-shift errors.
ACCESSION
BC039017.1 GI:24659234
VERSION
BC039017
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 2972)
AUTHORS

ALIGNMENTS

BC039017 2972 bp mRNA linear HTC 19-NOV-2003
Homo sapiens meningoeloma expressed antigen 6 (coiled-coil
proline-rich), mRNA (cDNA clone IMAGE:5770231), containing
frame-shift errors.
BC039017
BC039017.1 GI:24659234
Homo sapiens (human)
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2972)
Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonald, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,

AK029466 Mus muscu
AY409373 Mus muscu
BC018567 Mus muscu
BC029513 Homo sapi
BX328783 BX328783
BX391393 BX391393
BX563897 AGENCOURT
BX376997 BX376997
BX328066 BX328066
BX400140 BX400140
BX376998 BX376998
CB852737 UI-CF-FNO
BX377913 BX377913
BX111137 BX111137
EG719803 602691209
BX326021 BX326021
EG289351 602387471
EG740463 602633857
BU687930 UI-CF-EC1
BX391394 BX391394
BX390661 BX390661
AUI35481 AUI35481
BU942235 AGENCOURT
BI561045 603254248
BE781656 601470531
BI461930 603204268
BI462084 603205263
BF980666 602304385
BX368123 BX368123
AK053010 Mus muscu
BG117369 603346374
BU280057 AGENCOURT
BQ278480 AGENCOURT
BI829591 603079328
BX367676 BX367676
AIO82250 ox79h01.x
BI761502 603046640
CA414131 UI-H-E20-
BM977364 UI-CF-EN1
BF794269 602255469
AA931824 om83a04.s

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, J., Smalhus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC Help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Guittard, G., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masidillo, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantrispob, S., Thomas, P.J., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H., and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAX Plate: 82, Row: 1 Column: 9
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5174560
This clone has the following problem: frame shifted.

FEATURES

Location/Qualifiers
1..2972
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon.9606"
/clone="IMAGE:5770231"
/tissue_type="Brain, fetal, whole pooled"
/clone_lib="NIH MGC_121"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"

ORIGIN

Query Match 83.9%; Score 2654; DB 11; Length 2972;
Best Local Similarity 95.6%; Pred. No. 0;
Matches 2765; Conservative 0; Mismatches 84; Indels 42; Gaps 3;
QY 310 GGCCAGTGTACTCGGTGACCGCAGAGAGCCTCGACGCTATGGAGGAGCCTGGTGTCT 369
Db 52 GGCCGGGGTACTGTGGCACCACAGAGAGCAGCTTGGCGCTATGGAGAGCGCGGGCT 111
QY 370 ACCCTCAGCCCTACCTGGGGCTGGTCTGTGAGAGAGCTACGAGAGTGTGGCAGCACTA 429
Db 112 ACCCTCAACCGTATTGGGGCTGCTCCTGGAGAGCTACGAGGGTGTGGCAGCACTG 171
QY 430 CCTGAGGTATGAGACGAGATGAGAACTCTTATGTTTCCATCGGAAGCTGGTGTATGT 489
Db 172 CCTGAGGTATGAGACGAGATCTTAATCTTATGCTTTTCCATCGGAAGTGTGTATGT 231
QY 490 GCAGCTGTATTGGATTTTGTGTTCTCTCTTTTGTGGAGAGCTTTAGATCGGTT 549

Db 232 GCAGCTGTGTTGGATTTTGTGTTCTCTCTTTTGTGGAGAGTTTTAGATCGGTT 291
QY 550 AGGAGTCGGCTTTTACGTGGAGAGAGCAAAACACTTGGTGCACCGCTTTCTGGACTAAT 609
Db 292 AGAGTCGGCTTTTATGTGGAGAGAGAGAAAGCTTGTCTTAATGCTTTCTGGACTAAT 351
QY 610 GAAGAAATATGTAACCTACTTGAATAATTTAGCTTATTTCAAAAGAGTATGAGGCTAT 869
Db 352 GAAGAAATATGTAACCTACTTGAATAATTTAGCTTATTTCAAAAGAGTATGAGGCTAT 411
QY 670 GAAGTAGAGTCATCTTTTAGAGGATGCGCAGCTTTGAGAAGGAGGAGCAGAGAGCAG 729
Db 412 GAAGTAGAGTCATCTTTTAAGAGATGCGCAGCTTTGAGAAGGAGGCAAC--AGAAGCACA 468
QY 730 AGTTGAGGCAACCTGTGAAAAGCTGAACAGCTCAATTCGATTCGAATTCGAGGATGAATC 789
Db 469 AGTTGAGGCAACCTGTGAAAAGCTGAACAGCTCAATTCGATTCGAATTCGAGGATGAATC 528
QY 790 CTCTGCTAGAAAAGAGCTTTAAAACCAAGAGAAATTAACCAATTTCTCAACAGATGAATG 849
Db 529 CTCTGCTAGAAAAGAGCTTTAAAAGAGAGAAATTCACACATTTCTGAACAAGATGAATG 588
QY 850 ATGCGGATATTTCAAAAAGTATACAGTCTCTAGAGAGATGAGTCAAAATCCCTCAATCA 909
Db 589 ATGCGGATATTTCAAAAAGGATACAGTCTCTAGAGAGATGAGTCAAAATCCCTCAATCA 648
QY 910 CAATAGCTGAGGCCAAATCATCTGCAAGACATTTAAAATGATGAGTGAAGAACGACGGCT 969
Db 649 CAATAGCTGAGGCCAAATATGACCTTCAAGATTTTCAATGATGAGTGAAGAACGATGAAG 708
QY 970 ATAGCAATAAAGATGCTTTTGAATGAAAATCTCAACTTCAGACAGCCATTAACAGCTT 1029
Db 709 ATAGCAATAAAGATGCTTTTGAATGAAAATCTCAACTTCAGAGAAAGCCAGAAACAGCTT 768
QY 1030 TTTGAGCAAGAGCTGAGTATGAGAGGAGAGTGAAGTGAATTAATAACAGAGAAATA 1089
Db 769 TT---GCAAGAGCTGAGTATGAGAGAGAGAGTGAAGTGAATTAATAACAGAGAGTA 825
QY 1090 ACATTTGAAGACTCCAAAGTACACGCGAGAACAACTTCTGAATGATAAAGAAAATCAATC 1149
Db 826 ACATTTGAAGACTCCAAAGTACATGCGAGAACAACTTCTAAATGATAAAGAAAGTCAATC 885
QY 1150 AAGACCTGATGACACTTCCCAATGATGAAGATCAGCTGCTGCTTGAAGAGAGAC 1209
Db 886 AAGACTCTGACTGAACGCTTGTAAAGATGAAGAGTGGGCTGCTTGTGGAGAGAC 945
QY 1210 ACAACGATGATGATAACCTGGAATTAGAAGTGAACAGTGAATCGGAAAATGCTGTAC 1269
Db 946 ATAAGGATGATGATACTTGAATTAGAATGAACAGTGAATCGGAAAATGCTGTAC 1005
QY 1270 TTA-----GATATCTCTCCAAAGGAGCT 1293
Db 1006 TTAGGTATTAAGTCACTCATCTCTCTTTTGTGTAATGATAATCTCTCAAAAGGAGCT 1065
QY 1294 TTGAAGAACTGATTCATGCTGCTAAAGTAAATGCTTTTAAACACCTTAGAAGAGAA 1353
Db 1066 TTGAAGAACTGATTCATGCTGCTAAAGTAAATGCTTTTAAACACCTTAGAAGAGAA 1125
QY 1354 AGAAACCAATTTATATTCAGTGTCTGAAGTGAATAAACAAGGAGAGCTTACAGAG 1413
Db 1126 AGAAACCAATTTATATTCAGTGTCTGAAGTGAATAAACAAGGAGAGCTTACAGAG 1185
QY 1414 CATATTAATAATCTTCAGACTCAACAGCACTTTTGCAGTCAGAAAAACACATTTTGA 1473
Db 1186 CATATTAATAATCTTCAGACTCAACAGCACTTTTGCAGTCAGAAAAACACATTTTGA 1245
QY 1474 AATGAGATCAGAGCTTCAACAGAACTTAAAGTATGACTGAATTAATCAAGAAAT 1533
Db 1246 AATGAGATCAGAGCTTCAACAGAACTTAAAGTATGACTGAATTAATCAAGAAAT 1305
QY 1534 GAATCAACCTCCAGGAAATTAACAGTGAAGGAAATTTATCGTTAGAGAGAGAGAG 1593
Db 1306 GAATCAACCTCCAGGAAATTAACAGTGAAGGAAATTTATCGTTAGAGAGAGAG 1365

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2004, 01:51:06 ; Search time 12060 Seconds
(without alignments)
11371.257 Million cell updates/sec

Title: US-09-744-314-14
Perfect score: 3164
Sequence: 1 aggttattcatgaagaag.....agtgatttaaaaaaaaaa 3164

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sta: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sta: *
28: em_un: *
29: em_vi: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pln: *
35: em_htg_rod: *
36: em_htg_man: *
37: em_htg_vrt: *
38: em_sy: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
C 1	2832	89.5	59130	9	AL135904	AL135904 Human DNA
C 2	2832	89.5	130030	9	AC004889	AC004889 Homo sapi
C 3	2832	89.5	140915	9	AC005587	AC005587 Homo sapi
C 4	2830.4	89.5	130030	9	AC004889	AC004889 Homo sapi
C 5	2804.2	88.6	157801	2	AC144412	AC144412 Homo sapi
C 6	2804.2	88.6	169723	2	AC106862	AC106862 Homo sapi
C 7	2804.2	88.6	175228	2	AC143343	AC143343 Homo sapi
C 8	2804.2	88.6	200123	9	AC073264	AC073264 Homo sapi
C 9	2764.2	87.4	3676	9	HSU34780	U94780 Human menin
C 10	2749.8	86.9	162320	2	AC076959	AC076959 Homo sapi
C 11	2603.6	82.3	202634	2	AC147075	AC147075 Pan trogl
C 12	2593.8	82.0	2931	9	HSM807156	EX640994 Homo sapi
C 13	2583	81.6	2957	9	HSM808160	EX648014 Homo sapi
C 14	2531.8	80.0	2890	9	HSM803540	AL832233 Homo sapi
C 15	2531.4	80.0	2895	9	BC064355	BC064355 Homo sapi
C 16	2503.8	79.1	158310	9	AL162571	AL162571 Human DNA
C 17	2496.4	78.9	150942	9	AC017070	AC017070 Homo sapi
C 18	2490.2	77.1	199636	2	AP001126	AP001126 Homo sapi
C 19	2439.4	77.1	200606	9	AC099849	AC099849 Homo sapi
C 20	2439.4	76.6	178151	9	AL136123	AL136123 Human DNA
C 21	2422.2	76.6	349980	6	AX711880	AX711880 Sequence
C 22	2422.2	76.6	161790	2	AC022371	AC022371 Homo sapi
C 23	2415	76.3	161863	9	AL606970	AL606970 Human DNA
C 24	2415	76.3	161863	9	AX335827	AX335827 Sequence
C 25	2414.6	76.3	2814	6	AX410773	AX410773 Sequence
C 26	2414.6	76.3	2814	6	AX454750	AX454750 Sequence
C 27	2414.6	76.3	2814	6	AX491228	AX491228 Sequence
C 28	2414.6	76.3	2814	6	AX658161	AX658161 Sequence
C 29	2414.6	76.3	2814	6	HSU32682	U73682 Human menin
C 30	2414.6	76.2	2865	9	BC038527	BC038527 Homo sapi
C 31	2410.6	76.2	169104	9	AL162377	AL162377 Human DNA
C 32	2403.2	75.2	2778	9	AF338233	AF338233 Homo sapi
C 33	2378	75.2	2778	9	BC043153	BC043153 Homo sapi
C 34	2340.2	74.0	2620	6	AX113851	AX113851 Sequence
C 35	2326.2	73.5	4345	9	AF273058	AF273058 Homo sapi
C 36	2326.2	73.5	4345	9	AL355585	AL355585 Human DNA
C 37	2302.2	72.8	76321	9	AC140149	AC140149 Homo sapi
C 38	2302.2	72.8	179950	2	AL162716	AL162716 Human DNA
C 39	2294.6	72.5	122916	9	HSM801357	BC051363 Homo sapi
C 40	2261.2	71.5	2664	9	BC051363	BC051363 Homo sapi
C 41	2247.4	71.0	2547	9	BC051363	AL158816 Homo sapi
C 42	2225.2	70.3	22885	2	AL158816	AF338234 Homo sapi
C 43	2207.2	69.8	4110	9	AF338234	AF338234 Homo sapi
C 44	2182.4	69.0	2651	9	HSM804728	AL833415 Homo sapi
C 45	2177.4	68.8	2601	9	BC031065	BC031065 Homo sapi

ALIGNMENTS

RESULT 1
AL135904/c
LOCUS
DEFINITION

59130 bp DNA linear PRI 20-JUL-2000
Human DNA sequence from clone RP5-1005H11 on chromosome 6 Contains
part of the gene for a 7 transmembrane receptor (rhodopsin family)
(olfactory receptor like) protein, the gene for WUSC44-D10988615.3
protein, part of the PDPN3 (phosphodiesterase I/nucleotide
pyrophosphatase 3) gene, STSs and GSSs, complete sequence.

ACCESSION AL135904
VERSION AL135904.11 GI:7159399
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Pred. No. is the number of results predicted by chance to have a

/note="match: GSS: Em:AQ627093"
complement(24780..25383)
/note="match: GSS: Em:AQ627184"
26053..26548
/note="TIGER1 repeat: matches 1..521 of consensus"
26549..26861
/note="LIM4 repeat: matches 4025..4339 of consensus"
26865..27121
/note="TIGER3b repeat: matches 2..272 of consensus"
27122..27424
/note="AluX repeat: matches 1..302 of consensus"
27425..27579
/note="TIGER3b repeat: matches 272..451 of consensus"
27582..27658
/note="TIGER3b repeat: matches 1155..1231 of consensus"
27659..29067
/note="LIM4 repeat: matches 2585..4035 of consensus"
29189..29287
/note="LIM4 repeat: matches 2499..2600 of consensus"
29316..29499
/note="match: STS: Em:AA931824"
29507..29990
/note="match: STS: Em:AA931824"
29530..29981
/note="match: GSS: Em:AQ356766"
29530..29737
/note="match: GSS: Em:AQ253729"
29533..29855
/note="match: GSS: Em:AQ319835"
29538..29943
/note="match: STS: Em:AA235111"
29560..29787
/note="match: STS: Em:G27651 Em:T99793"
complement(29779..32112)
/gene="dJ1005H11.2"
/gene="dJ1005H11.2"
/note="match: proteins: Tr:O95046"
/codon_start=1
/evidence=not experimental
/product="dJ1005H11.2 (WUGSC:H_DJ0988G15.3 PROTEIN)"
Query Match 89.5%; Score 2832; DB 9; Length 59130;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 2991; Conservative 0; Mismatches 160; Indels 13; Gaps 5;
QY 1 AGCTTTAATCATGAAGACACGACATTTTAAAGTGTTATTCACCAAAAATAAAGCTT 60
Db 32463 AGCTTTAATCATGAAGACACGACATTTTAAAGTGTTATTCACCAAAAATAAAGCTT 32404
QY 61 CAAATATGTGATGTAAGAACTGCCAGAACTAAGCGCGCGCGCTCAGACCCAGCGCTGC 120
Db 32403 CAAATATGTGATGTAAGAACTGCCAGAACTAAGCGCGCGCGCTCAGACCCAGCGCTGC 32344
QY 121 CTCAGGATGTAAGTGTAAACAAGGSCCAGGGAGGTGTGGGGACACATGGGCGCTG 180
Db 32343 CTCAGGATGTAAGTGTAAACAAGGSCCAGGGAGGTGTGGGGACACATGGGCGCTG 32284
QY 181 TGAGGCTGTGGGTGCGCGGTTCCCGAGTCCCGCCGAGCCCGCTCCACAGTGTCCG 240
Db 32283 TGAGGCTGTGGGTGCGCGGTTCCCGAGTCCCGCCGAGCCCGCTCCACAGTGTCCG 32224
QY 241 CTCGGTGTGTTCAGTGGCGGATTCGGGTTCAGACCCAGGCTCGGTTCACACC 300
Db 32223 CTCGGTGTGTTCAGTGGCGGATTCGGGTTCAGACCCAGGCTCGGTTCACACC 32164
QY 301 GCTTGTGTGGCCAGTGTACTCGGTGACCGCCAGAGCAGCTCCGACGCTATGGAGGAG 360
Db 32163 GCTTGTGTGGCCAGTGTACTCGGTGACCGCCAGAGCAGCTCCGACGCTATGGAGGAG 32104
QY 361 CCGTGTGTACCCCTCAGCCCTACTCGGGGTGTCTCGAGGAGCTACCGAGTTGTG 420
Db 32103 CCGTGTGTACCCCTCAGCCCTACTCGGGGTGTCTCGAGGAGCTACCGAGTTGTG 32044

QY 421 CGACCACTACCTGAGAGTATGAGACCAAGATCCTTATGGTTTTCATCGAACTG 480
Db 32043 CGACCACTACCTGAGAGTATGAGACCAAGATCCTTATGGTTTTCATCGAACTG 31984
QY 481 GTGTAATGTGAGCTGTTATTTGGATTTTGTGTTCTCCTTTTGTGGAGAGTTT 540
Db 31983 GTGTAATGTGAGCTGTTATTTGGATTTTGTGTTCTCCTTTTGTGGAGAGTTT 31924
QY 541 AGATCGGTTAGAGTCCGCTTTAGCTGGGAAGAGACAAAACTTGTGTCAACCTTTCT 600
Db 31923 AGATCGGTTAGAGTCCGCTTTATGTGGGAAGAGACAAAACTTGTGTCAACCTTTCT 31864
QY 601 GGACTAATTTGAAGAAAATGTAACTACTTTGAAAAATTTAGCCTTATTTCAAAAAGATAT 660
Db 31863 GGACTAATTTGAAGAAAATGTAACTACTTTGAAAAATTTAGCCTTATTTCAAAAAGATAT 31804
QY 661 GAAGCTATGAAGTAGAGTCACTCTTTAGAGATCCAGCTTTTGAGAGGAGGCGAGCAGAA 720
Db 31803 GAAGCTATGAAGTAGAGTCACTCTTTAGAGATCCAGCTTTTGAGAGGAGGCGAGCAGAA 31744
QY 721 GAAGCAGAACTTTGGAGGCAACCTGTGAAAAGCTGAACAGGTCCAAATTTCTGAACCTTGAG 780
Db 31743 GAAGCAGAACTTTGGAGGCAACCTGTGAAAAGCTGAACAGGTCCAAATTTCTGAACCTTGAG 31684
QY 781 GATGAATCCCTCTGCTAGAAAAGAGCTTAAAAAAGAGATCTAAACATTTCTCAACAA 840
Db 31683 GATGAATCCCTCTGCTAGAAAAGAGCTTAAAAAAGAGATCTAAACATTTCTCAACAA 31624
QY 841 GATGAATTTGATGGCGGATATTTCAAAAAGTATACAGTCTCTAGAAAGTAGAGTCAAAATCC 900
Db 31623 GATGAATTTGATGGCGGATATTTCAAAAAGTATACAGTCTCTAGAAAGTAGAGTCAAAATCC 31564
QY 901 CTCAAATCACAATAGCTGAAGCCAAATCATCTGCAAGACATTTAAATCAGGTGAAGAA 960
Db 31563 CTCAAATCACAATAGCTGAAGCCAAATCATCTGCAAGACATTTAAATCAGGTGAAGAA 31504
QY 961 CGACGGCTATAGCAATAAAGAGATGCTTTGAATGAAAATTTCTCAACTTCAGCAAGCCAT 1020
Db 31503 CGACGGCTATAGCAATAAAGAGATGCTTTGAATGAAAATTTCTCAACTTCAGCAAGCCAT 31444
QY 1021 AAAAGCTTTTTCAGCAAGAGCTGAAGTATGGAAGAGAGAGTGAAGTGAATTAATAA 1080
Db 31443 AAAAGCTTTTTCAGCAAGAGCTGAAGTATGGAAGAGAGAGTGAAGTGAATTAATAA 31384
QY 1081 CAGAAAAATAATTTGAAGCTCCAAAGTACACGACAGAACTTCTGAATGATAAGAA 1140
Db 31383 CAGAAAAATAATTTGAAGCTCCAAAGTACACGACAGAACTTCTGAATGATAAGAA 31324
QY 1141 AATCAATCAAGAACCTTGAATGATGGAACATTTGAAAGTGAAGTGAAGTGAAGTGAAGT 1200
Db 31323 AATCAATCAAGAACCTTGAATGATGGAACATTTGAAAGTGAAGTGAAGTGAAGTGAAGT 31264
QY 1201 GAAGAGACACAAAGGATGATGAATGAACTGGAATTTAGAAAGTGAACAGTGAATCGAAAT 1260
Db 31263 GAAGAGACACAAAGGATGATGAATGAACTGGAATTTAGAAAGTGAACAGTGAATCGAAAT 31204
QY 1261 GGTGCTTACTTAGATAATCTCCAAAAGAGCTTTGAAAGAACTGATTCATGCTGCTAAG 1320
Db 31203 GGTGCTTACTTAGATAATCTCCAAAAGAGCTTTGAAAGAACTGATTCATGCTGCTAAG 31144
QY 1321 TTAATGCTTCTTTAAACCTTTAGAGGAGAAAGAACCAATTTATATTCAGTTGCT 1380
Db 31143 TTAATGCTTCTTTAAACCTTTAGAGGAGAAAGAACCAATTTATATTCAGTTGCT 31084
QY 1381 GAAGTGTATAAACAAGAGAGCTTACAGAGCATATTTAAAAATTTTCAAGACTCAACAA 1440
Db 31083 GAAGTGTATAAACAAGAGAGCTTACAGAGCATATTTAAAAATTTTCAAGACTCAACAA 31024
QY 1441 GCATCTTTCAGTCAAGAAACACACATTTTGAATAGAGTCAAGAGCTTCAACAGAAA 1500
Db 31023 GCATCTTTCAGTCAAGAAACACACATTTTGAATAGAGTCAAGAGCTTCAACAGAAA 30964

QY	1501	CTTTAAAGTAAATGACTGAATTTATATCAAGAAAATGAAATCGAACTCCACAGGAAATTAACA	1560
Db	30963	CTTTAAATATATGACTGAATTTCTATCAAGAAAATGAAATCGAACTCTCAGGAAATTAACA	30904
QY	1561	GTAGAGGAAAAATATCGGTTAGAGAAAAGAGAACTTTCTTAAAGTAGATGAAAAAGATC	1620
Db	30903	GTGAGGAAAAATTCACGAATAGAGGAAGAGAGAGCTTTCTAGAGTGAAGAAAAGATC	30844
QY	1621	AGCCATGCCACTGAAGAGCTGGAGACCTATAGAAAGCGAGGCCAAAGATCTTTGAAGAAGAA	1680
Db	30843	AGCCATGCCACTGAAGAGCTGGAGACCTATAGAAAGCTAGCCAAAGATCTTTGAAGAAGAA	30784
QY	1681	TTGGAGAGAACTATTCAATCTTTATCAAGGCGCAGATTATTTCCTCATGAGAAAAAGCACAT	1740
Db	30783	TTGGAGAGAACTGTTTCATTTTTATCAAAAGCAGGTTATTTCCTACGAGAAAAGAGGACAT	30724
QY	1741	GATAATTTGGTGGCAGCTCGGAATGCTGAAGGAACCTCAATGATTTAAGGAAAGAAAT	1800
Db	30723	GATAATTTGGTGGCAGCTCGGAATGCTGAAGGAACCTCAGTGAATTAAGGAAAGAAAT	30664
QY	1801	GCTCACAAACAGACAAAAATTAACCTGAAAACAGAGCTTAAATTTGAACTTTTAGAAAAAGAT	1860
Db	30663	GCTCACAAACAAACAAAAATTAACCTGAAAAGAGAGTTGAAATTTGAACTTTTAGAAAAAGAT	30604
QY	1861	CCATTATGCACTGCATGTTCCCAATACAGCATTTGGCAGAGAGCAATCCCATATATGGTCCC	1920
Db	30603	CCATTATGCACTGCATGTTTCAATACAGCATTTGGCAGAGAGCAATCCCATATGATGCC	30544
QY	1921	TCACCAATGGTGTGGCTTCATCTGAAAACAGAGCTTTTCTCTCTCCTCCAACTTTGTTG	1980
Db	30543	TCACCAATGGGTGGGCTTCATCTGAAAACAGAGCTTTTCCCTCTCTCCAACTTTGTTG	30484
QY	1981	GAGGGTCCACTCAGACTCTCACCTTTGCTTCCAGGGGAGGAGGAGGCTCACAGGSC	2040
Db	30483	GAGGATCCACTCAGACTCTCACCTGTGCTTCCAGGGGAGGAGGAGGCCCAAGCAGC	30424
QY	2041	CCAGGGAATCCTCTGGACCAATCAGATTACCAATGAAAAGAGGAGAAATCAAGCTGTGATAGG	2100
Db	30423	CCAGGGAATCCCCTGGACCAATCAGATTACCAATGAAAAGAGGAGAACCAAGCTATGACAGG	30364
QY	2101	TTAACCGATCCTCATAGGGCTCCCTCTGACACTGGGTCCTCTGTCACTCCATGGGACACAG	2160
Db	30363	TTAATCGATCCTCACAGGGCTCCTTCAGACATGGGTCCTCTGTCACTCCGCTGGAAACAG	30304
QY	2161	GACCGTAGGATGATTTCTCCGCGCAGGACCAATCATATCCTGTATTAGCCCTTCCTCCA	2220
Db	30303	GACCGTAGGATGATTTCTCCACACAGGCAATCATATCCTGTATTCAACTCTCTCCTCCA	30244
QY	2221	CAAAAGCAACAGAGATTTGCTCTTAATCTGTGTAGACTGTCTGGACACGACAGACTCAGA	2280
Db	30243	CAAAAGGGAACAGAGATTTTAATTTAATTCGAAAGACTGTCTGGACACGACAGAACCCAGA	30184
QY	2281	AGTTTTAATATGCCCTTTTGGATAAAATGGATGGGTCAATGCCCTTCAGAAATGGAATCC	2340
Db	30183	AGTTTTAAAATGACTTTCTTTGGATAAAATGGATAGTCAATGGCTTCAGAAATGGAATCC	30124
QY	2341	AGTAGAAATGATACCAAGATGATCTTTGTAATTTAAATGTGCTGATTCATCTCTCCCT	2400
Db	30123	AGTAGAAATGATGCCAAGATGATCTTTGTAATTTAAATGTGCTGATTCATCTCTCCCT	30064
QY	2401	GCTGAAAATGAAGCACTGGCCCTGGCTTTGTTCCCTCCACCTTTTGCTCCAATTCAGAGGT	2460
Db	30063	GCTGAAAATGAAGCAATGGCCCTGGCCCTTATTTCTCCACCTCTTGCTCCAATTCAGCGGT	30004
QY	2461	CCATTGTTCCAGTGGATGCAAGAGGGCCCATTTCTGAGAGAGGACCTCTCTTCCCCCCA	2520
Db	30003	CCATTGTTCCAGTGGATACAGGGGCCCATTCATGAGAGAGGACCTCTCTTCCCCCCA	29944
QY	2521	CCTCCTCCAGAGCCATGTTTGGAGCTTCTCGAGATTAATTTCCCAAGAGGATTTCCA	2580
Db	29943	CCTCCTCCAGSAACCATGTTTGGAGCTTCTCGAGGTTAATTTTCCCAAGAGGATTTCCA	29884
QY	2581	GCTCCACACACTGCTCCATTGCAATGAGAAAATGTCTATCCACCGAGGGGTTTCTCCTCT	2640

[illegible]

RESULT 2
AC004889
LOCUS
DEFINITION Homo sapiens PAC clone RP4-798C17 from 7, complete sequence.
ACCESSION AC004889
KEYWORDS HYG.
VERSION AC004889.1 GI:4156187
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 130030)
AUTHORS Sulston, J.E. and Wilson, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 5 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 130030)
AUTHORS Gibson, A., and Kozlowski, Z.
TITLE The sequence of Homo sapiens PAC clone RP4-798C17
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 130030)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-1998) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 130030)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (14-JAN-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

```

7841. 8087
/rpt family="Li"
8098. 8672
/rpt family="Li"
10048. 10070
/rpt family="AT_rich"
10071. 10201
/rpt family="Li"
10472. 10595
/rpt family="GA-rich"
10597. 10824
/rpt family=" (GGAA)n"
11168. 11193
/rpt family="AT_rich"
11194. 11732
/rpt family="Li"
11733. 12143
/rpt family="MaLR"
12143. 13177
/rpt family="Li"
13178. 13549
/rpt family="MaLR"
13550. 15665
/rpt family="Li"
15665. 15929
/rpt family="Rcksha"
15928. 17402
/rpt family="Rcksha"
17428. 17608
/rpt family="Li"
17610. 17974
/rpt family="Li"
18051. 18316
/rpt family="Li"
18317. 18619
/rpt family="Alu"
18620. 18858
/rpt family="Li"
18869. 19137
/rpt family="ERV_L"
19241. 19353
/rpt family="ERV_L"
19354. 19879
/rpt family="MER1_type"
19880. 19999
/rpt family="ERV_L"
21348. 21680
/rpt family="Li"
21709. 22051
/rpt family="MaLR"
23392. 23933
/rpt family="MaLR"
25951. 26252
/rpt family="Alu"
26503. 26680
/rpt family="Alu"
27013. 27185
/rpt family="MaLR"
27186. 27549
/rpt family="MaLR"
27350. 27672
/rpt family="MaLR"
27920. 28076
/rpt family="ERV_L"
31894. 32190
/rpt family="MER2_type"
32191. 32335
/rpt family="Alu"
32361. 32441
/rpt family="MER2_type"
32609. 32670
/rpt family="AT_rich"
32699. 32819

```

```

/rpt_family="L1"
repeat_region 32820..33201
/rpt_family="MaLR"
repeat_region 33202..34039
/rpt_family="L1"
repeat_region 34104..34866
/rpt_family="L1"
repeat_region 37906..37943
/rpt_family="AT_rich"
repeat_region 37945..38043
/rpt_family="L1"
38044..38148
/product="snRNA-U6-related"
repeat_region 38044..38148
/rpt_family="U6"
repeat_region 38165..38548
/rpt_family="L1"
repeat_region 40180..40317
/rpt_family="L2"
repeat_region 40707..41010
/rpt_family="L2"
repeat_region 41410..41718
/rpt_family="Alu"

Query Match 89.5%; Score 2832; DB 9; Length 130030;
Best Local Similarity 94.5%; Pred No. 0;
Matches 2991; Conservative 0; Mismatches 160; Indels 13; Gaps 5;

QY 1 AGGTTTAATCCATGAAGAAGACAGCAATTTTAAAGTGTATTACCAAAAATAAAGCTT 60
DB 34768 AGGTTTAATCCATGAAGAAGACAGCAATTTTAAAGTGTATTACCAAAAATAAAGCTT 34827
QY 61 CAAAATATCTGATGCAAACTCCAGAACTAAAGCGGGCGGGCTCAGACAGCGCTGC 120
DB 34828 CAAAATATCTGATGCAAACTCCAGAACTAAAGCGGGCGGGCTCAGACAGCGCTGC 34887
QY 121 CTCAGGATGAAGTGTAAACAAGAGGCGCAGGGAGGTGGTGGGGGCAACAATGGGCGTG 180
DB 34888 CTCAGGATGAAGTGTAAACAAGAGGCGCAGGGAGGTGGTGGGGGCAACAATGGGCGTG 34947
QY 181 TGAGGCGCTGGGTGGCGCGCGTCCCGAGCTCCCGCGAGCCCGCTCCACAGTGGTCCG 240
DB 34948 TGAGGCGCTGGGTGGCGCGCGTCCCGAGCTCCCGCGAGCCCGCTCCACAGTGGTCCG 35007
QY 241 CTCGGTGTGGTGTCTACGTGCGCATTCGGGTTCAGACCCCAAGGCTGGTGTTCCTACC 300
DB 35008 CTCGGTGTGGTGTCTACGTGCGCATTCGGGTTCAGACCCCAAGGCTGGTGTTCCTACC 35067
QY 301 GCTTGTGTGGCCAGTGTACTGCGGTGACCCGAGCAGCGCTCGACGCTATGGAGGAG 360
DB 35068 GCTTGTGTGGCCAGTGTACTGCGGTGACCCGAGCAGCGCTCGACGCTATGGAGGAG 35127
QY 361 CCTGTGTCTACCCCTCAGCCCTACCTGGGCTGGTCTGGAGGAGCTACGCGAGTGTG 420
DB 35128 CCTGTGTCTACCCCTCAGCCCTACCTGGGCTGGTCTGGAGGAGCTACGCGAGTGTG 35187
QY 421 GCAGCACTACTGAGAGTATGAGACAGATGAGAACTCTATGGTTTTCATCGGAACTG 480
DB 35188 GCAGCACTACTGAGAGTATGAGACAGATGAGAACTCTATGGTTTTCATCGGAACTG 35247
QY 481 GTGTGTGTGCGAGCTGTATTGATGATTTTGTGTCTCTCTTTTGTGGAGAGTGT 540
DB 35248 GTGTGTGTGCGAGCTGTATTGATGATTTTGTGTCTCTCTTTTGTGGAGAGTGT 35307
QY 541 AGATCGGTTAGGAGTCGGCTTTTACGTGGGAAGAGAGCAAAAACCTTGGTGCACCGTTCT 600
DB 35308 AGATCGGTTAGGAGTCGGCTTTTACGTGGGAAGAGAGCAAAAACCTTGGTGCACCGTTCT 35367
QY 601 GGACTAATGAAGAAAATGTAACACTACTTGAATAATTTAGCCCTATTCAAAAAGAGTAT 660
DB 35368 GGACTAATGAAGAAAATGTAACACTACTTGAATAATTTAGCCCTATTCAAAAAGAGTAT 35427
QY 661 GAAGGCTATGAAGTAGAGTCACTCTTTAGAGGATGCCAGCTTTGAGAAGGAGGAGCAGAA 720

```

```

DB 35428 GAAGGCTATGAAGTAGAGTCACTCTTTAGAGGATGCCACTTTGAGAAGCGCGCAGCAA 35487
QY 721 GAAGCAGCAAGTTTGGAGGCAACCTGTGAAAAGCTGAAACAGCTGCAATTTCTGAACTTGAG 780
DB 35488 GAAGCAGCAAGTTTGGAGGCAACCTGTGAAAAGCTGAAACAGCTGCAATTTCTGAACTTGAG 35547
QY 781 GATGAAATCCTCTGTCTAGAAAAGACTTAAACAAGAGAAATCTTAAACATTTCTCAACAA 840
DB 35548 GATGAAATCCTCTGTCTAGAAAAGACTTAAACAAGAGAAATCTTAAACATTTCTCAACAA 35607
QY 841 GATGAAATCCTGTGCGGATATTTCAAAAAGTATACAGTCTCTAGAAAAGTGAAGTCAAAATCC 900
DB 35608 GATGAAATCCTGTGCGGATATTTCAAAAAGTATACAGTCTCTAGAAAAGTGAAGTCAAAATCC 35667
QY 901 CTCAAATCAAAATAGCTGAAGCCAAAATCATCTGCAAGACATTTTAAATAGTGAAGAA 960
DB 35668 CTCAAATCAAAATAGCTGAAGCCAAAATCATCTGCAAGACATTTTAAATAGTGAAGAA 35727
QY 961 CGACGGGCTATAGCAATAAAGATGCTTTGAATGAAAATTTCTCAACTTCAGACAAGCCAT 1020
DB 35728 CGACGGGCTATAGCAATAAAGATGCTTTGAATGAAAATTTCTCAACTTCAGACAAGCCAT 35787
QY 1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGGAAGGAGAGTGAAGTGAATTAATAA 1080
DB 35788 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGGAAGGAGAGTGAAGTGAATTAATAA 35847
QY 1081 CAGAAAATAACATTTTGAAGACTCCAAAGTACAGCGAGAAACAAGTTCTGAATGATAAGAA 1140
DB 35848 CAGAAAATAACATTTTGAAGACTCCAAAGTACAGCGAGAAACAAGTTCTGAATGATAAGAA 35907
QY 1141 AATCACATCAAGACCTGTGACACTTGCCTGCAATGATGAAGATCAGGCTGCTGTGCTT 1200
DB 35908 AATCACATCAAGACCTGTGACACTTGCCTGCAATGATGAAGATCAGGCTGCTGTGCTT 35967
QY 1201 GAAGAGACACACAGGATGATGATTAACCTGGATTAAGAGTGAACAGTGAATCGGAAAAAT 1260
DB 35968 GAAGAGACACACAGGATGATGATTAACCTGGATTAAGAGTGAACAGTGAATCGGAAAAAT 36027
QY 1261 GGTGCTTACTTAGATTAATCCTCCAAAGGAGCTTTGAAGAACTGATTCATGCTGCTAAG 1320
DB 36028 GGTGCTTACTTAGATTAATCCTCCAAAGGAGCTTTGAAGAACTGATTCATGCTGCTAAG 36087
QY 1321 TTAATGCTTTCTTTAAACCTTTAGAGGAGAAAGAAACCAATTTATATTCAGTTGTCT 1380
DB 36088 TTAATGCTTTCTTTAAACCTTTAGAGGAGAAAGAAACCAATTTATATTCAGTTGTCT 36147
QY 1381 GAAGTTGATTAACAAAGGAGAGCTTACAGAGCATATTAAATAATCTTCAGACTCAACAA 1440
DB 36148 GAAGTTGATTAACAAAGGAGAGCTTACAGAGCATATTAAATAATCTTCAGACTCAACAA 36207
QY 1441 GCATCTTTGAGTCAAGAAAACACACATTTTGAATAAGAGATCAGAGCTTCACAGAAA 1500
DB 36208 GCATCTTTGAGTCAAGAAAACACACATTTTGAATAAGAGATCAGAGCTTCACAGAAA 36267
QY 1501 CTTAAAGTAAATGACTGAATTTATATCAAGAAAATGAATGAATCCACAGGAAATTAACA 1560
DB 36268 CTTAAAGTAAATGACTGAATTTATATCAAGAAAATGAATGAATCCACAGGAAATTAACA 36327
QY 1561 GTAGAGAAAATTTATCGGTTAGAGAAAAGAGAACTTTCTAAAGTGAATGAAGATC 1620
DB 36328 GTAGAGAAAATTTATCGGTTAGAGAAAAGAGAACTTTCTAAAGTGAATGAAGATC 36387
QY 1621 AGCCATGCACTGAAGAGCTGGAGACCTATAGAAAAGCAGTACCAAGATCTTGAAGAA 1680
DB 36388 AGCCATGCACTGAAGAGCTGGAGACCTATAGAAAAGCAGTACCAAGATCTTGAAGAA 36447
QY 1681 TTGAGAGAACTATTTCATCTTATCAAGGGGAGATTTATTTCCNCTGAGAAAAGACAT 1740
DB 36448 TTGAGAGAACTATTTCATCTTATCAAGGGGAGATTTATTTCCNCTGAGAAAAGACAT 36507
QY 1741 GATTAATCGTGGAGCTCGCAATTCGTAAGAAACCTCAATGATTTTAAAGAAAGAAAT 1800

```

Db 36508 GATAATTGGTGGCAGCTCGGACTGCTGAAAGAAACCTCAGTGATTTAAGGAAAGAAAT 36567
QY 1801 GCTCACAACAGACAAAATTAACCTGAAACAGAGCTTAAATTTGAACTTTTAGAAAAAGAT 1860
Db 36568 GCTCACAACAGACAAAATTAACCTGAAACAGAGCTTAAATTTGAACTTTTAGAAAAAGAT 36627
QY 1861 CCTATGCACTGATGTTCCAAATACAGATTTGGCAGAGAGCATCCCATATGGTCC 1920
Db 36628 CCTAATGCACTGATGTTCCAAATACAGATTTGGCAGAGAGCATCCCATATGGTCC 36687
QY 1921 TCACCAATGGGTGGGCTTCATCTGAAACAGAGCTTTCTCTCTCCTCCCAACTTTGTTG 1980
Db 36688 TCACCAATGGGTGGGCTTCATCTGAAACAGAGCTTTCTCTCTCCTCCCAACTTTGTTG 36747
QY 1981 GAGGCTCCTACAGATCTCACTCTGCTTCCAGGGGGAGAGAGGCTTCAGAGGC 2040
Db 36748 GAGGATCCACTCAGATCTCACTCTGCTTCCAGGGGGAGAGAGGCTTCAGAGGC 36807
QY 2041 CCAGGGAATCCTCTGACCATCAGATACCAATGAAAGAGAGATCAAGCTGTGATAGG 2100
Db 36808 CCAGGGAATCCTCTGACCATCAGATACCAATGAAAGAGAGATCAAGCTGTGATAGG 36867
QY 2101 TTAACCGATCTCATAGGCTCCTCTGACACTGGCTCTCTGACCTCCATGGGACAG 2160
Db 36868 TTAATCGATCTCATAGGCTCCTCTGACACTGGCTCCTCTGACCTCCATGGGACAG 36927
QY 2161 GACCTAGGATGATGTTCTCTCCAGGAGACATCATATCCTGATTCAGCCCTTCTCCA 2220
Db 36928 GACCTAGGATGATGTTCTCTCCAGGAGACATCATATCCTGATTCAGCCCTTCTCCA 36987
QY 2221 CAAAGGCAAGACAGATTTGTTCTTAATCTGTGATGATCTCTGACAGCAGAACTCAGA 2280
Db 36988 CAAAGGCAAGACAGATTTGTTCTTAATCTGTGATGATCTCTGACAGCAGAACTCAGA 37047
QY 2281 AGTTTAAATGCTCTTTGGTAAATGGATGGCTCAATGCTTCAGAAATGGAATCC 2340
Db 37048 AGTTTAAATGCTCTTTGGTAAATGGATGGCTCAATGCTTCAGAAATGGAATCC 37107
QY 2341 AGTAGAAATGATACCAAGATGATCTTTGTAATTTAAATGTGCTGATTCATCTCTCC 2400
Db 37108 AGTAGAAATGATACCAAGATGATCTTTGTAATTTAAATGTGCTGATTCATCTCTCC 37167
QY 2401 GCTGAAATGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
Db 37168 GCTGAAATGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 37227
QY 2461 CCATTGTTCCAGTGAATGCAAGAGGCCATCTTCAGAGAGAGACCTCTTCCGCCCA 2520
Db 37228 CCATTGTTCCAGTGAATGCAAGAGGCCATCTTCAGAGAGAGACCTCTTCCGCCCA 37287
QY 2521 CTCCTCCAGAGCAATGTTTGGAGCTTCGAGATATTTTCCACCAAGGATTTCCCA 2580
Db 37288 CTCCTCCAGAGCAATGTTTGGAGCTTCGAGATATTTTCCACCAAGGATTTCCCA 37347
QY 2581 GGTCCACACCTGCTCCATTTGCAATGAGAAATGCTATCCACCGAGGGTTTCTCTCT 2640
Db 37348 GGTCCACACCTGCTCCATTTGCAATGAGAAATGCTATCCACCGAGGGTTTCTCTCT 37407
QY 2641 TACCTTCCCAAGACCTGATTTTCCGCCACCCACATTCAGAGTGAAGTGA 2700
Db 37408 TACCTTCCCAAGACCTGATTTTCCGCCACCCACATTCAGAGTGAAGTGA 37465
QY 2701 TTCCCTCAGGTTGATTCACCTTCAATGAGCCTGCTACTGAAATCCAGACACAG 2760
Db 37466 TTCCCTCAGGTTGATTCACCTTCAATGAGCCTGCTACTGAAATCCAGACACAG 37525
QY 2761 CAAGAAACCTGACAAATTTTCTCTTCAAAAGTAAATTTGACTGATCTCATTTTCA 2820
Db 37526 CAAGAAACCTGACAAATTTTCTCTTCAAAAGTAAATTTGACTGATCTCATTTTCA 37585
QY 2821 GTTTAAGTAACTGTTTACTTAAGTGAATTTGACTTTTGTCTCAATTTGAAGCTTAATGA 2880
Db 37586 GTTTAAGTAACTGTTTACTTAAGTGAATTTGACTTTTCTCAATTTGAAGCTTAATGA 37644

QY 2881 ATTATAATCTCAGGATGATTTTGTAAATAAAGATGATTTAAATATGAATCTTATGAG 2940
Db 37645 ATAATAGTTCTCAGGATGATTTTGTAAATAAAGATGATTTAAATATGAATCTTATGAG 37704
QY 2941 TAATTAATTTCAATTTT-----TATTTAGAGGATATAACTATTTCAAT--TTGATTAATC 2993
Db 37705 TAATTAATTTCCATTTTATTTATTTCTAGATCATATAACTTTTAACTTGGTGAATTAATC 37764
QY 2994 CACTATTATATAACAATAGTGGAGTTTATATATGTAATCTTTCAGTGGGAGGCTT 3053
Db 37765 CACTTTAGAGAAACAATAGTGGAGTTTATATATGTAATCTTGCAGTGGAGGCTT 37824
QY 3054 TAATTTCTGAAGTCT---GTGCTTTATGCAAGACTGATTTTACTGTGTTGTGACA 3110
Db 37825 TAAATCTTAAAGTGTGGTGTCTTCATGCAAGACTGATTTTACTGTGTTGTAGATA 37884
QY 3111 AATGTGAAGTAACTTTATGCTTAAATAAATTTATAGTTGATTTA 3154
Db 37885 AATGTGAAGTAACTTTATGCTTAAATAAATTTATAGTTGATTTA 37928

RESULT 3
AC005587 140915 bp DNA linear PRI 07-NOV-2002
LOCUS Homo sapiens PAC clone RP5-988G15 from 6, complete sequence.
DEFINITION AC005587
ACCESSION AC005587
VERSION AC005587.1 GI:4156166
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 140915)
Sulston, J.E. and Waterston, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
9847074
REFERENCE 2 (bases 1 to 140915)
Hou, S., Wohldmann, P. and Le, T.
The sequence of Homo sapiens PAC clone RP5-988G15
Unpublished (2001)
REFERENCE 3 (bases 1 to 140915)
Waterston, R.H.
Direct Submission
Submitted (01-SEP-1998) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 140915)
Waterston, R.H.
Direct Submission
Submitted (14-JAN-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5 (bases 1 to 140915)
Waterston, R.
Direct Submission
Submitted (15-JAN-1999) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 6 (bases 1 to 140915)
Waterston, R.
Direct Submission
Submitted (07-NOV-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jan 14, 1999 this sequence version replaced gi:3907499.
----- Genome Center
Center: Washington University Genome Sequencing Center

Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@atson.wustl.edu
----- Summary Statistics -----
Center project name: H_DJ0988G15

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/csc>

SOURCE INFORMATION:
This clone was derived from human PAC library RPL-5, prepared by Pieter de Jong and coworkers at <http://www.chori.org> using the method described by Iannone et al., Nature Genetics 6:84-9 (1994). The library is from one male donor.
The clone may be obtained either from Genome Systems, Inc. (<http://www.genomesystems.com>) or Research Genetics, Inc. (<http://www.resgen.com>); or from Pieter de Jong.
VECTOR: pCYPAC2

NEIGHBORING SEQUENCE INFORMATION:
Actual start of this clone is at base position 1 of AC005587;
actual end is at 140915 of AC005587.

FEATURES
SOURCE

source	1..140915	/organism="Homo sapiens"	/mol_type="genomic DNA"	/db_xref="taxon:9606"	/chromosome="6"	/map="e"	/clone="RPS-988G15"	/clone_lib="RPC1-5"	13..2090	/rpt_family="L1"	2097..2239	/rpt_family="L1"	2300..2603	/rpt_family="Alu"	2604..4120	/rpt_family="L1"	4123..4166	/rpt_family="(TA)n"	4168..4307	/rpt_family="L1"	6635..6668	/rpt_family="AT-rich"	6825..7260	/rpt_family="L1"	7261..7642	/rpt_family="MatR"	7643..7923	/rpt_family="L1"	7987..8283	/rpt_family="Alu"	8286..8412	/rpt_family="L1"	8413..8708	repeat_region
--------	-----------	--------------------------	-------------------------	-----------------------	-----------------	----------	---------------------	---------------------	----------	------------------	------------	------------------	------------	-------------------	------------	------------------	------------	---------------------	------------	------------------	------------	-----------------------	------------	------------------	------------	--------------------	------------	------------------	------------	-------------------	------------	------------------	------------	---------------

repeat_region	/rpt_family="Alu"	8709_..8838	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="Li"	8840_..8886	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="MaLR"	9111_..9139	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="AT_rich"	9204_..9490	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="Alu"	9846_..9924	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="Alu"	10286_..10455	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="L2"	10759_..10838	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="Mariner"	11127_..11266	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="L2"	11850_..12154	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="Alu"	12172_..12211	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="AT_rich"	12352_..12469	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="(TACG)n"	13458_..12607	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="(CATA)n"	12560_..12733	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="(TA)n"	12660_..12777	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="(CATA)n"	12778_..13031	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="Alu"	13226_..13535	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="L2"	15273_..15520	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="MIR"	15954_..16022	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="MaLR"	16023_..16332	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="MBR2_type"	16333_..16371	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="MaLR"	16630_..16742	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="MaLR"	16816_..17068	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="Li"	17088_..17348	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="Li"	17857_..17979	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="Alu"	18014_..18506	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="MaLR"	18707_..19081	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="MaLR"	19135_..20717	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="Retroviral"	20728_..20748	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="AT_rich"	21444_..21599	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
repeat_region	/rpt_family="MIR"	21852_..22330	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
misc_feature	/note="similar to Mus musculus EST C88518 (NID:g3164280)"	21852_..22254	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
misc_feature	/note="similar to Mus musculus EST W13172 (NID:g1287301)"	21852_..22254	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
misc_feature	/note="similar to Mus musculus EST W23487 (NID:g1300321)"	21852_..22127	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
misc_feature	/note="similar to Mus musculus EST W10529 (NID:g1284855)"	21852_..22127	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
misc_feature	/note="similar to EST AA082271 (NTD:g16424330)"	21857_..22052	to EST AA082271 (NTD:g16424330)	zn37g03.rl"
misc_feature	/note="similar to EST AA082271 (NTD:g16424330)"	21857_..22052	to EST AA082271 (NTD:g16424330)	zn37g03.rl"

```
misc_feature 21663..22255
/notes="similar to Mus musculus EST AA274895 (NID:g1914893)
va92c10.r1"
misc_feature 21665..22255
/notes="similar to EST AI097080 (NID:g3446662) oz22c05.x1"

Query Match      89.5%; Score 2832; DB 9; Length 140915;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 2991; Conservative 0; Mismatches 160; Indels 13; Gaps 5;

QY 1 AGGTTTAAATCCATGAAGAAGACGACATTTTAAAGTGATTCACCAAAATAAAGCTT 60
DB 91632 AGGTTTAAATCCATGAAGAAGACGACATTTTAAAGTGATTCACCAAAATAAAGCTT 91691
QY 61 CAAATAATGTGATGTGAAATCTGCCAGAACTAAAGCGCGCGCTCAGACCCAGCGCTGC 120
DB 91692 CAAATAATGTGATGTGAAATCTGCCAGAACTAAAGCGCGCGCTCAGACCCAGCGCTGC 91751
QY 121 CTGAGGATGTAAAGTGTAACAAGAGGCCAGGGAGGTGTGGGGACACATGGGCCCTG 180
DB 91752 CTGAGGATGTAAAGTGTAACAAGAGGCCAGGGAGGTGTGGGGACACATGGGCCCTG 91811
QY 181 TGAGGCGCTGTGGGTGCCCGGCTTCCCGAGCTCCCGCGAGCCCGCTCCACAGTGGTCCG 240
DB 91812 TGAGGCGCTGTGGGTGCCCGGCTTCCCGAGCTCCCGCGAGCCCGCTCCACAGTGGTCCG 91871
QY 241 CTCGGTGTGTGTACGTGCGGATTCGGGTTCAGACCCCAAGCTGCGTGTCTCCACC 300
DB 91872 CTCGGTGTGTGTACGTGCGGATTCGGGTTCAGACCCCAAGCTGCGTGTCTCCACC 91931
QY 301 GCTTGTGTGGCCAGTGTACTGCGGTGACCGCCAGAGCAGCTCCAGCGTATGGAGGAG 360
DB 91932 GCTTGTGTGGCCAGTGTACTGCGGTGACCGCCAGAGCAGCTCCAGCGTATGGAGGAG 91991
QY 361 CTGGTGTCTACCTCAGCCCTTACCTGGGCTGTCTGGAGAGCTACCGAGTGTG 420
DB 91992 CTGGTGTCTACCTCAGCCCTTACCTGGGCTGTCTGGAGAGCTACCGAGTGTG 92051
QY 421 GCAGCACTACTCTGAGGTATGAGACCAAGATGAGATCTTATGTTTCCATCGGAACCTG 480
DB 92052 GCAGCACTACTCTGAGGTATGAGACCAAGATGAGATCTTATGTTTCCATCGGAACCTG 92111
QY 481 GTGATATGTGAGCTGTATTTGATTTTGTGTTCTCCTTTTGTGGAGAGTGT 540
DB 92112 GTGATATGTGAGCTGTATTTGATTTTGTGTTCTCCTTTTGTGGAGAGTGT 92171
QY 541 AGATCGGTGTAGAGTCCGCTTTACCTGGAGAGAGCAAAACCTGGTGCAACGCTTCT 600
DB 92172 AGATCGGTGTAGAGTCCGCTTTATGTGGAGAGAGCAAAACCTGGTGCAACGCTTCT 92231
QY 601 GGACTAATTTGAAGAAAATGTAACTACTTGAATAATTTAGCCCTTATTTCAAAAGAGTAT 660
DB 92232 GGACTAATTTGAAGAAAATGTAACTACTTGAATAATTTAGCCCTTATTTCAAAAGAGTAT 92291
QY 661 GAAGGCTATGAAGTAGAGTCACTTTAGAGGATCCAGCTTGAGAGGAGGAGGAGGAGAA 720
DB 92292 GAAGGCTATGAAGTAGAGTCACTTTAGAGGATCCAGCTTTGAGAGGAGGAGGAGGAGAA 92351
QY 721 GAAGCAGGAATTTGGAGGCAACCTGTGAAAGCTGAACAGGTCCAAATTTCTGAACCTTGAG 780
DB 92352 GAAGCAGGAATTTGGAGGCAACCTGTGAAAGCTGAACAGGTCCAAATTTCTGAACCTTGAG 92411
QY 781 GATGAAATCTCTGTCTAGAAAAGAGCTTAAACAAGAGAAATCTAAACATTTCTCAACAA 840
DB 92412 GATGAAATCTCTGTCTAGAAAAGAGCTTAAACAAGAGAAATCTAAACATTTCTCAACAA 92471
QY 841 GATGAAATGTAGGGGGATATTTCAAAAGATATACAGTCTCTAGAGATGAGTCAAAATCC 900
DB 92472 GATGAAATGTAGGGGGATATTTCAAAAGATATACAGTCTCTAGAGATGAGTCAAAATCC 92531
QY 901 CTCAAATCACAATAATAGCTGAAGCCMAATATCTGCAAGACATTTAAATGAGTGAAGAA 960
DB 92532 CTCAAATCACAATAATAGCTGAAGCCMAATATCTGCAAGACATTTAAATGAGTGAAGAA 92591
```

```
QY 961 CGACGGGCTATAGCAATAAAGAGTGTCTTTGAATGAAAATTTCTCAACTTCACACAAGCCAT 1020
DB 92592 CGACGGGCTATAGCAATAAAGAGTGTCTTTGAATGAAAATTTCTCAACTTCACACAAGCCAT 92651
QY 1021 AAACAGCTTTTTCAGCAAGAAGAGCTGAAGTATGGAAGAGAGAGTGAAGTAAATAA 1080
DB 92652 AAACAGCTTTTTCAGCAAGAAGAGCTGAAGTATGGAAGAGAGAGTGAAGTAAATAA 92711
QY 1081 CAGAAAAATAACATTTTGAAGAGTCCAAAGTACACGCAGAACCAAGTTCTCAATGATAAGAA 1140
DB 92712 CAGAAAAATAACATTTTGAAGAGTCCAAAGTACACGCAGAACCAAGTTCTCAATGATAAGAA 92771
QY 1141 AATCATCTCAAGACCCCTGACTGGACACCTTGCCAATGATGAAAGATCAGGCTGTGTGCTT 1200
DB 92772 AATCATCTCAAGACCCCTGACTGGACACCTTGCCAATGATGAAAGATCAGGCTGTGTGCTT 92831
QY 1201 GAAGAAGACACAAACCGATGATGAATCTGGAATTTAGAAGTGAAACAGTGAATCGAAAAAT 1260
DB 92832 GAAGAAGACACAAACCGATGATGAATCTGGAATTTAGAAGTGAAACAGTGAATCGAAAAAT 92891
QY 1261 GGTGCTTACTTAGATAAATCTCCAAAAGAGGCTTTGAAGAACTGATTCATGCTGTGAAG 1320
DB 92892 GGTGCTTACTTAGATAAATCTCCAAAAGAGGCTTTGAAGAACTGATTCATGCTGTGAAG 92951
QY 1321 TTAAATGCTTCTTTAAACCTTTAGAGGAGAGAGAACCAATTTATTTCAAGTGTCT 1380
DB 92952 TTAAATGCTTCTTTAAACCTTTAGAGGAGAGAGAACCAATTTATTTCAAGTGTCT 93011
QY 1381 GAAGTTGATAAACAAGAGAGAGCTTACAGAGCATATTTAAAAATCTTTCAGACTCAACAA 1440
DB 93012 GAAGTTGATAAACAAGAGAGAGCTTACAGAGCATATTTAAAAATCTTTCAGACTCAACAA 93071
QY 1441 GCATCTTTGCGAGTCAAGAACACACATTTTGAAGTGAAGTCAAGAGCTTCAACAGAAA 1500
DB 93072 GCATCTTTGCGAGTCAAGAACACACATTTTGAAGTGAAGTCAAGAGCTTCAACAGAAA 93131
QY 1501 CTTAAAGTAAAGACTGAATTTATATCAAGAAAATGAAATGAAATCTTCAAGGAAATTAACA 1560
DB 93132 CTTAAAGTAAAGACTGAATTTATATCAAGAAAATGAAATGAAATCTTCAAGGAAATTAACA 93191
QY 1561 FTAGAGGAAAATTTATCGGTAGAGAAAGAGAACTTTCTAAAGTGAAGTGAAGAGATC 1620
DB 93192 FTAGAGGAAAATTTATCGGTAGAGAAAGAGAACTTTCTAGAGTGAAGAGAAAGATC 93251
QY 1621 AGCCATGCCACTGAAGAGCTGGAGACCTTATAGAAAAGCGAGGCAAGATCTTGAAGAGAA 1680
DB 93252 AGCCATGCCACTGAAGAGCTGGAGACCTTATAGAAAAGCTAGCCAAAGATCTTGAAGAGAA 93311
QY 1681 TTGGAGAGAACTATTTCAATTTATCAAGGCGAGATTTTCCCATGAGAAAAGACAT 1740
DB 93312 TTGGAGAGAACTATTTCAATTTATCAAGGCGAGATTTTCCCATGAGAAAAGAGACAT 93371
QY 1741 GATAATTTGGTTGGCAGCTCGGAATCTGAAAAGAAACCTCAATGATTTAAAGAAAAGAAAAT 1800
DB 93372 GATAATTTGGTTGGCAGCTCGGAATCTGAAAAGAAACCTCAATGATTTAAAGAAAAGAAAAT 93431
QY 1801 GCTCAACACAGCAAAATAATTAATCTGAAAACAGAGCTTAAATTTGAACTTTTGAAGAAAAGAT 1860
DB 93432 GCTCAACACAGCAAAATAATTAATCTGAAAAGAGAGTTGAAATTTGAACTTTTGAAGAAAAGAT 93491
QY 1861 CCTTATGCACTCGATGTTTCCAAATACAGATTTGSCAGAGAGCATTTCCCATATCGTCCC 1920
DB 93492 CCTTATGCACTCGATGTTTCCAAATACAGATTTGSCAGAGAGCATTTCCCATATCGTCCC 93551
QY 1921 TCACCATTTGGTGGCCTTCAATCTGAAAACAGAGCTTTTCTCTCTCTCCCTCAACTTTGTG 1980
DB 93552 TCACCATTTGGTGGCCTTCAATCTGAAAACAGAGAGCTTTTCTCTCTCTCTCAACTTTGTG 93611
QY 1981 GAGGCTCCACTCAGACTCTCAGCTTTGCTTCCAGGGGAGGAGGAGGCTCAGAGGC 2040
DB 93612 GAGGCTCCACTCAGACTCTCAGCTTTGCTTCCAGGGGAGGAGGAGGCTCAGAGGC 93671
```


QY	2041	CMAGGGAATCTCTGGA	CAATCAGATTACCAATGAAGAGGAGAAATCAAGCTGTGATAGG	2100
Db	93672	CCAGGGAATCCCTCGAC	CAATCAGATTACCAATGAAGAGGAGAAACCAAGCTATGACAGG	93731
QY	2101	TTAAACCGATCTCTCAT	AGGGCTCCCTCTGACACTGGGTCTGTCTCACTCCATCCATGGAACGAC	2160
Db	93732	TTAAATCGATCCTCAC	AGGGCTCCTTCTGACACTGGGTCCCTGTCACTCCTCGGTGGAACAG	93791
QY	2161	GACCGTAGATGATGTTT	CTCTCGGCAGGACAAATCATATCTGTGATTGAGCCCTTCCTCCA	2220
Db	93792	GACCGTAGATGATGTTT	CCTCCAACAGGGCAATCATATCTGTGATTCAACTCTTCCTCCA	93851
QY	2221	CAAAGGCAAGACAGATT	TTGTTCTTAATCTCGTAGACTGTCTGGACCAAGCAAACTCAGA	2280
Db	93852	CAAAGGCAAGACAGATT	TTATTTCTAAATCTGAAGAGACTGTCTGGACCAAGCAAACTCAGA	93911
QY	2281	AGTTTAAATATGCTCT	TTTGATATAAATGGATGGTCAATGCCTTCAGAAATGGAATCC	2340
Db	93912	AGTTTAAATATGACTT	TTTGATATAAATGGATAGTCAATGCCTTCAGAAATGGAATCC	93971
QY	2341	AGTAGAAATGATACCA	AAGATGATCTTGTAATTTAAATGTGCTGATTCATCTCTCCCT	2400
Db	93972	AGTAGAAATGATGCC	AAAGATGATCTTGTAATTTAAATGTGCTGATTCATCTCTCCCT	94031
QY	2401	GCTGAAATCAGACCA	CTCGCTGGCTTGTTCCTCCACCTCTGTGCTCCAACTCAGAGGT	2460
Db	94032	GCTGAAATCAGACCA	CTCGCTGGCTTGTTCCTCCACCTCTGTGCTCCAACTCAGAGGT	94091
QY	2461	CCATGTTTCCAGTGG	ATCCAGAGCCCATCTTTGAGAAGGAGACCTCTTTTCCGCCCA	2520
Db	94092	CCATGTTTCCAGTGG	ATCCAGAGGAGCCCATCTGAGAAGGAGACCTCTTTTCCGCCCA	94151
QY	2521	CCCTCTCCAGGAGC	CAATGTTTGAGCTCTCCAGATTAATTTTCCACCAAGGGAATTCCTCA	2580
Db	94152	CCCTCTCCAGGAA	CAATGTTTGAGCTCTCCAGGTTAATTTTCCACCAAGGGAATTCCTCA	94211
QY	2581	GGTCCACCACTGCT	CCATTTGCAATGAGAAATGTCTATCCACGAGGGTTCCTCTCT	2640
Db	94212	GGTCCACCACTGCT	CCATTTGCAATGAGAAACATCTATCCACGAGGGTTCCTCTCT	94271
QY	2641	TACCTTCCCCCAAG	ACTGGATTTTTCCTCCCAACCCACATCTCTGAAAGGTAGAAAGTGAG	2700
Db	94272	TACCTTCACTCGA	GAAGCTGGATTTT--ACCCTCAACCCACATCTCTGAAAGGTAGAAAGTGAG	94329
QY	2701	TTCCCTCAGGTTTG	ATTCACCTTCAAAATGAGCTGCTACTGAAACATCCAGAACCCACAG	2760
Db	94330	TTCCCTCAGGATTG	ATTCGCTTCAAAAGGAGCTGCTACTGGAACATCCAGAACCCACAG	94389
QY	2761	CAAGAAACCTCGA	CAATATTTTCTCTCTTCAAAAGTAAATTTGACTGATCTCAATTTCA	2820
Db	94390	CAAGAACCTCGACA	ATATTTGCTTCTTCAAAAGTAAATTTGACTGATCTCAATTTCA	94449
QY	2821	GTTTAAAGTAACT	CTGTTTAAAGTGATTAACATTTTGTCTCAAAATGGAAGCTTAATGGA	2880
Db	94450	GTTTAAAGTAACT	CTGTTTAAAGTGATTAATGATTAATTTTCTCAAAATGGAAGCTTAATGGA	94508
QY	2881	ATTATTAATCTCAG	GATAGTATTTGTAATAAAGATGATTTAAATATGAATCTTATGAG	2940
Db	94509	ATAATAGTTCTCAG	GATAGTATTTGTAATAAAGATGATTTGTAATATGAATCTTATGAG	94568
QY	2941	TAAATTAATTTCA	ATTTT--TATTTTAGCGGTATACTATTTTCAAT--TTGATTAATC	2993
Db	94569	TAAATCAATTTTCC	ATTTTATTTATTTATATCTTAGATCAATACTTTAACTTGGTGAACATACT	94628
QY	2994	CACATTAATATAA	CAACTAGTGGAGTTTATATATGTAATCTTTTCAAGTGGGAGGCTT	3053
Db	94629	CACCTTTAGAGA	AAACAATAGTGGAGTTTATATATGTAATCTTTTCAAGTGGGAGGCTT	94688
QY	3054	TAAATTTCTGAAG	CTCTTATGCCAGAAGCTGATTTACTGTGTTGGACA	3110
Db	94689	TAAATTTCTGAAG	CTCTTATGCCAGAAGCTGATTTACTGTGTTGGACA	94748
QY	3111	AATGTGAAGTAAG	TAAGCTTATGCTTAAATTAATTAAGTTAGTTGATTTA	3154

Db 94749 AATGTGAAGTAACCTTTATGCTTAAATTAATAAATTTTCAATGA 94792

RESULT 4
AC004889/c
LOCUS AC004889 130030 bp DNA linear PRI 02-OCT-2003
DEFINITION Homo sapiens PAC clone RP4-798C17 from 7, complete sequence.
AC004889
ACCESSION AC004889.1 GI:4156187
VERSION AC004889.1
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 130030)
Sulston,J.E. and Wilson,R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
PUBMED 9847074
2 (bases 1 to 130030)
Gibson,A. and Kozlowski.
The sequence of Homo sapiens PAC clone RP4-798C17
Unpublished (2001)
3 (bases 1 to 130030)
Waterston,R.H.
Direct Submission
Submitted (12-JUN-1998) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 130030)
Waterston,R.H.
Direct Submission
Submitted (14-JAN-1999) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 130030)
Waterston,R.
Direct Submission
Submitted (26-APR-2003) Department of Genetics, Washington
University, 444 Forest Park Avenue, St. Louis, Missouri 63108, USA
6 (bases 1 to 130030)
Wilson,R.
Direct Submission
Submitted (02-OCT-2003) Department of Genetics, Washington
University, 444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jan 14, 1999 this sequence version replaced gi:3213108.
----- Genome Center
Center: Washington University Genome Sequencing Center
Web site: <http://genome.wustl.edu>
Contact: sapiens@watson.wustl.edu
----- Summary Statistics

Center project name: H_DJ0798C17

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
The sequence of this clone was established as part of a mapping and sequencing collaboration between the NEGR1 Chromosome 7 Mapping

Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GB/CHR7> ; send <mailto:edgreen@nhgri.nih.gov> , or see <http://genome.wustl.edu>

SOURCE INFORMATION:

This clone was derived from human PAC library RPCI-4, prepared by Pieter de Jong and coworkers at <http://www.chori.org> using the method described by Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from one male donor.

The clone may be obtained either from Genome Systems, Inc. (<http://www.genomesystems.com>) or Research Genetics, Inc. (<http://www.resgen.com>); or from Pieter de Jong

VECTOR:

FEATURES	source	Location/Qualifiers
		1..33030
		/organism="Homo sapiens"
		/mol_type="genomic DNA"
		/db_xref="taxon:9606"
		/chromosome="7"
		/map="7"
		/clone="RP4-798C17"
		/clone_lib="RPCI-4"
		5..3376
repeat_region		/rpt_family="L1"
repeat_region		3412..3820
repeat_region		/rpt_family="MER1_type"
repeat_region		4921..5004
repeat_region		/rpt_family="(TAGA)n"
repeat_region		5035..5081
repeat_region		/rpt_family="(TA)n"
repeat_region		7528..7696
repeat_region		/rpt_family="L1"
repeat_region		7697..7840
repeat_region		/rpt_family="(GAAA)n"
repeat_region		7841..8087
repeat_region		/rpt_family="L1"
repeat_region		8098..8672
repeat_region		/rpt_family="L1"
repeat_region		10048..10070
repeat_region		/rpt_family="AT_rich"
repeat_region		10071..10201
repeat_region		/rpt_family="L1"
repeat_region		10472..10595
repeat_region		/rpt_family="GA_rich"
repeat_region		10597..10624
repeat_region		/rpt_family="(GAAA)n"
repeat_region		11168..11193
repeat_region		/rpt_family="AT_rich"
repeat_region		11194..11732
repeat_region		/rpt_family="L1"
repeat_region		11733..12143
repeat_region		/rpt_family="MALR"
repeat_region		12144..13177
repeat_region		/rpt_family="L1"
repeat_region		13178..13549
repeat_region		/rpt_family="MALR"
repeat_region		13550..15665
repeat_region		/rpt_family="L1"
repeat_region		15666..15929
repeat_region		/rpt_family="RICKsha"
repeat_region		15928..17402
repeat_region		/rpt_family="RICKsha"
repeat_region		17428..17608
repeat_region		/rpt_family="L1"
repeat_region		17810..17974
repeat_region		/rpt_family="L1"
repeat_region		18051..18316
repeat_region		/rpt_family="L1"
repeat_region		18317..18619
repeat_region		/rpt_family="Alu"

repeat_region	18620..18958 /rpt_family="L1"
repeat_region	18869..19137 /rpt_family="ERVL"
repeat_region	19241..19353 /rpt_family="ERVL"
repeat_region	19354..19879 /rpt_family="MER1_type"
repeat_region	19880..19999 /rpt_family="ERVL"
repeat_region	21248..21680 /rpt_family="L1"
repeat_region	21709..22051 /rpt_family="MaLR"
repeat_region	23592..23933 /rpt_family="MaLR"
repeat_region	25951..26252 /rpt_family="Alu"
repeat_region	26503..26680 /rpt_family="Alu"
repeat_region	27013..27185 /rpt_family="MaLR"
repeat_region	27186..27549 /rpt_family="MaLR"
repeat_region	27550..27672 /rpt_family="MaLR"
repeat_region	27920..28076 /rpt_family="ERVL"
repeat_region	31894..32190 /rpt_family="MER2_type"
repeat_region	32191..32335 /rpt_family="Alu"
repeat_region	32361..32441 /rpt_family="MER2_type"
repeat_region	32609..32670 /rpt_family="AT_rich"
repeat_region	32699..32819 /rpt_family="L1"
repeat_region	32820..33201 /rpt_family="MaLR"
repeat_region	33202..34039 /rpt_family="L1"
repeat_region	34104..34866 /rpt_family="L1"
repeat_region	37905..37943 /rpt_family="AT_rich"
repeat_region	37945..38043 /rpt_family="L1"
snRNA	38044..38148 /product="snRNA-U6-related"
repeat_region	38044..38148 /rpt_family="U6"
repeat_region	38165..38548 /rpt_family="L1"
repeat_region	40180..40317 /rpt_family="L2"
repeat_region	40707..41010 /rpt_family="L2"
repeat_region	41410..41718 /rpt_family="Alu"

Query Match 89.5%; Score 2830.4; DB 9; Length 130030;
Best Local Similarity 94.5%; Pred.No.0;
Matches 2990; Conservative 0; Mismatches 161; Indels 13; Gaps 5;

py	1	AGGTTTAACTCCATGAAGAAGACAGCAAAATTTTAAAAAGTGTAATTCACCACAAAATAAAGCTT	60
bb	121216	AGGTTTAACTCCATGAAGAAGACAGCAAAATTTTAAAAAGTGTAATTCACCACAAAATAAAGCTT	1211157
yy	61	CAAATATGTGATGTGAAACTGCCAGAACCTAAGCGGGCCGGGCTCAGACGAGCGCTGC	120
yy	121156	CAAATATGTGATGTGAAACTGCCAGAACCTAAGCGGGCCGGGCTCAGACGAGCGCTGC	121097

Query Match

Query Match	99.35%	Score 2050.47	EE 5;
Best Local Similarity	94.58%	Score No. 0;	
Matches 2930;	Conservative	0; Mismatches	161;
Indels	13;	Gaps	5;

Qy 1 AGGTTTAACTCCATGAAGAAGACAGCAATTTTAAAGTGATTCAACAAAAATTAAGCTT 60

nb 121215 AGGTTTAACTCCATGAAGAAGACAGCAATTTTAAAGTGATTCAACAAAAATTAAGCTT 121157

61 CAAATATGTGATGTGAAACTGCCAGAACTAAGCGCGCGGCTCAGACCCAGCGCTGC 120

D_b 121156 CAAAATATGTGATGTGAAAACTGCCAGAACTAAGCGGGCGGCTCAGACCAGCGCTG 121097

QY 121 CTCAGGATGTAAGTGTAAACAAGAGGCGCAGGAGGTGGTGGGGACAACTGGCGCTG 180
DB 121096 CTCAGGATGTAAGTGTAAACAAGAGGCGCAGGAGGTGGTGGGGACAACTGGCGCTG 121037
QY 181 TGAGGCGCTGTGGGTGCGCGCGTTCCTCCAGCTCCCGCGCAGCCCGCTCCACAGTGGTCCG 240
DB 121036 TGAGGCGCTGTGGGTGCGCGCGTTCCTCCAGCTCCCGCGCAGCCCGCTCCACAGTGGTCCG 120977
QY 241 CTCGGGTGTTGTGTCACGTGGCATTCGGGTTCAGAGCCCAAGGCTGGTGTCTCCACC 300
DB 120976 CTCGGGTGTTGTGTCACGTGGCATTCGGGTTCAGAGCCCAAGGCTGGTGTCTCCACC 120917
QY 301 GCTTGTGTCGCGAGTACTGCGGTGACCGCCAGAGCAGCCTCGAGCTATGAGGAG 360
DB 120916 GCTTGTGTCGCGAGTACTGCGGTGACCGCCAGAGCAGCCTCGAGCTATGAGGAG 120857
QY 361 CTTGTGTCTACCCCTCAGCCCTACCTGGGCTGTCTCTGGAGGAGCTACGAGAGTGTG 420
DB 120856 CTTGTGTCTACCCCTCAGCCCTACCTGGGCTGTCTCTGGAGGAGCTACGAGAGTGTG 120797
QY 421 GCAGCACTACCTGAGATGAGACAGATGAGATCCTTATGTTTTCCTCGGAACTG 480
DB 120796 GCAGCACTACCTGAGATGAGACAGATGAGATCCTTATGTTTTCCTCGGAACTG 120737
QY 481 GTGGTATGTCAGCTGTTATTTGGATTTTGTGTTCTCTCTTTTGTGGAGAGTTT 540
DB 120736 GTGGTATGTCAGCTGTTATTTGGATTTTGTGTTCTCTCTTTTGTGGAGAGTTT 120677
QY 541 AGATCGGTTAGAGTGGGCTTTACGTGGAGAGAGCAAACTTGGTGCACAGCTTCT 600
DB 120676 AGATCGGTTAGAGTGGGCTTTACGTGGAGAGAGCAAACTTGGTGCACAGCTTCT 120617
QY 601 GGACTAATTTGAAGAAATTTGTAATCTATTGTAATTTAGCTTATTTCAAAAGAGTAT 660
DB 120616 GGACTAATTTGAAGAAATTTGTAATCTATTGTAATTTAGCTTATTTCAAAAGAGTAT 120557
QY 661 GAAGGCTATGAGTAGAGTATCTTTAGAGGATGCCAGCTTTGAGAAGGAGGCGACAGAA 720
DB 120556 GAAGGCTATGAGTAGAGTATCTTTAGAGGATGCCAGCTTTGAGAAGGAGGCGACAGAA 120497
QY 721 GAAGCAAGAGTTGGAGGCACTCTGTAAGAGCTGAAAGCTGCAAGCTGCAATTTGAACTTGG 780
DB 120496 GAAGCAAGAGTTGGAGGCACTCTGTAAGAGCTGAAAGCTGCAAGCTGCAATTTGAACTTGG 120437
QY 781 GATGAATCTCTGTCTAGAAAGAGCTTAAACAGAGCAATCTAAACATTTCTCAACAA 840
DB 120436 GATGAATCTCTGTCTAGAAAGAGCTTAAACAGAGCAATCTAAACATTTCTCAACAA 120377
QY 841 GATGAATTTGATGGCGGATATTTCAAAAGTATACAGTCTCTAGAAAGAGTCAAAATCC 900
DB 120376 GATGAATTTGATGGCGGATATTTCAAAAGTATACAGTCTCTAGAAAGAGTCAAAATCC 120317
QY 901 CTCAATCACAATAGCTGAGCAAAATCATCTGCAAGACATTTAAATGAGTGAAGAA 960
DB 120316 CTCAATCACAATAGCTGAGCAAAATCATCTGCAAGACATTTAAATGAGTGAAGAA 120257
QY 961 CGACGGGCTATAGCAATAAAGATGCTTTGAATGAAATTTCTCACTTCAGACAAAGCCAT 1020
DB 120256 CGACGGGCTATAGCAATAAAGATGCTTTGAATGAAATTTCTCACTTCAGACAAAGCCAT 120197
QY 1021 AAACAGCTTTTTCAGCAAGAGTGAAGTGAAGAGGAGAGTGAAGTGAATTAATAA 1080
DB 120196 AAACAGCTTTTTCAGCAAGAGTGAAGTGAAGAGGAGAGTGAAGTGAATTAATAA 120137
QY 1081 CAGAAATTAACATTTGAAGTCTCAAAAGTACAGCGCAACAACTTCTGAATGATAAGAA 1140
DB 120136 CAGAAATTAACATTTGAAGTCTCAAAAGTACAGCGCAACAACTTCTGAATGATAAGAA 120077
QY 1141 AATCACATCAAGACCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1200
DB 120076 AATCACATCAAGACCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 120017
QY 1201 GAAGAGACACACAGGATGATGATTAACCTGGAATTTAGAGTGAAGTGAAGTGAAGT 1260

DB 120016 GAAGAGACACACAGGATGATGATTAACCTGGAATTTAGAAAGTGAACAGTCAATGGAAAT 119957
QY 1261 GGTGCTTACTTAGATTAATCTCTCAAAAGAGCTTTGAAGAACTGATTCATGCTGCTAAG 1320
DB 119956 GGTGCTTACTTAGATTAATCTCTCAAAAGAGCTTTGAAGAACTGATTCATGCTGCTAAG 119897
QY 1321 TTAATATGCTTCTTTTAAAAACCTTTAGAGGAGAAAGAAACCAAAATTTATTCAGTTGTCT 1380
DB 119896 TTAATATGCTTCTTTTAAAAACCTTTAGAGGAGAAAGAAACCAAAATTTATTCAGTTGTCT 119837
QY 1381 GAAGTTGATTAACAAAGAGAGCTTACAGAGCATATTAATAATCTTCAGACTCAACAA 1440
DB 119836 GAAGTTGATTAACAAAGAGAGCTTACAGAGCATATTAATAATCTTCAGACTCAACAA 119777
QY 1441 GCATCTTTTCAGTTCAGAAACACACATTTTGAATAATGAGATTCAGAGCTTCAACAGAAA 1500
DB 119776 GCATCTTTTCAGTTCAGAAACACATATTTTGAATAATGAGATTCAGAGCTTCAACAGAAA 119717
QY 1501 CTTAAAGTAAATGATGATTAATTCAGAGAAATGAAATGAAATCTTCAGAGAAATTAACA 1560
DB 119716 CTTAAAGTAAATGATGATTAATTCAGAGAAATGAAATGAAATCTTCAGAGAAATTAACA 119657
QY 1561 GTAGAGAAATTTATCGGTTAGAGAAAGAGAAACCTTTCTAAAGTAGATGAAAGATC 1620
DB 119656 GTAGAGAAATTTATCGGATAGAGAAAGAGAGAGCTTTCTAGAGTGAAGAAAGATC 119597
QY 1621 AGCATGCTCACTGAAGAGCTGAGACCTATAGAAAGAGGAGCAAGATCTTGAAGAGAA 1680
DB 119596 AGCATGCTCACTGAAGAGCTGAGACCTATAGAAAGAGGAGCAAGATCTTGAAGAGAA 119537
QY 1681 TTGAGAGAACTATTCATTTTATCAAGGCGAGATTTATTTCCCATGAGAAAAAGACAT 1740
DB 119536 TTGAGAGAACTATTCATTTTATCAAGGCGAGTATTTTCTACGAGAAAAAGAGACAT 119477
QY 1741 GATAATTTGGTGGCAGCTCGAAATGCTGAAAGAAACCTCAGTGATTTAAGGAAAGAAAT 1800
DB 119476 GATAATTTGGTGGCAGCTCGAAATGCTGAAAGAAACCTCAGTGATTTAAGGAAAGAAAT 119417
QY 1801 GCTCACAACAGACACAAATTTAATCTGAAACAGAGCTTTAAATTTGAACTTTTGAAGAAAGAT 1860
DB 119416 GCTCACAACAGACACAAATTTAATCTGAAACAGAGTTCGAATTTGAACTTTTGAAGAAAGAT 119357
QY 1861 CTTTATCCACTCGATGTTCCAAATACAGCATTTTGGCAGAGAGCATTTCCCATATGTTCC 1920
DB 119356 CTTTATCCACTCGATGTTCCAAATACAGCATTTTGGCAGAGAGCATTTCCCATATGTTCC 119297
QY 1921 TCACCATTTGGTGGCTTCTCATCTGAAACAGAGCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1980
DB 119296 TCACCATTTGGTGGCTTCTCATCTGAAACAGAGCTTTTCTCTCTCTCTCTCTCTCTCTCTCT 119237
QY 1981 GAGGCTCCACTCAGACTCTCACCTTTCTTCCAGGGGAGAGAGAGAGCTCACAGGC 2040
DB 119236 GAGGCTCCACTCAGACTCTCACCTTTCTTCCAGGGGAGAGAGAGAGCTCACAGGC 119177
QY 2041 CCAGGGAATCTCTGGACCATCAGATTTACCAATGAAGAGAGAGATCAAGCTGTGATAGG 2100
DB 119176 CCAGGGAATCTCTGGACCATCAGATTTACCAATGAAGAGAGAGATCAAGCTGTGATAGG 119117
QY 2101 TTAACCGATCTCATAGGGTCTCTCTGACATCTGGGTCTCTGTCACTCTCTCTCTCTCTCTCTCT 2160
DB 119116 TTAACCGATCTCATAGGGTCTCTCTGACATCTGGGTCTCTGTCACTCTCTCTCTCTCTCTCTCT 119057
QY 2161 GACGTAGGATGATTTTCTCTCCAGGAGCAATCATATCTGATTCAGCCCTTCTCTCTCA 2220
DB 119056 GACGTAGGATGATTTTCTCTCCAGGAGCAATCATATCTGATTCAGCCCTTCTCTCTCA 118997
QY 2221 CAAAGGCAAGACAGATTTTGTCTAAATCTGGTGAAGTGTCTGACAGAGAGTCTCAGA 2280
DB 118996 CAAAGGCAAGACAGATTTTGTCTAAATCTGGTGAAGTGTCTGACAGAGAGTCTCAGA 118937
QY 2281 AGTTTAAATATGCTTCTTTTGGATTAATAATGAGTGGGTCAATGCTTCTCAGAAATGGAATCC 2340

TITLE The sequence of Homo sapiens clone
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 157801)
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (12-APR-2003) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA

COMMENT

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 Contact: submissions@wustl.edu
 ----- Project Information -----
 Center project name: H_NH0224H01
 ----- Summary Statistics -----
 Sequencing vector: M13; #
 Chemistry: Dye-terminator Big Dye; # of reads
 Chemistry: Dye-terminator Big Dye; # of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 158459 bases at least Q40
 Consensus quality: 158679 bases at least Q30
 Consensus quality: 158958 bases at least Q20

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 27114: contig of 27114 bp in length
 * 27115 27214: Gap of unknown length
 * 27215 157801: contig of 130587 bp in length.

FEATURES
 source
 1..157801
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="UNK"
 /clone="RP11-224H1"
 1..27114
 /note="assembly_name:Contig13"
 27215..157801
 /note="assembly_name:Contig14"

ORIGIN
 Query Match 98.6%; Score 2804.2; DB 2; Length 157801;
 Best Local Similarity 94.2%; Pred. No. 0;
 Matches 2981; Conservative 0; Mismatches 163; Indels 19; Gaps 6;
 Qy 1 AGCTTTAATCCATGAGAGACAGCAATTTTAAAGTGTATTACCAAAAAATAAGCTT 60
 Db 79615 AGCTTTAATTCATGAAGAAGACAGCAATTTTAAAGTGTATTACCAAAAAATAAGCTT 79556
 Qy 61 CAAATAATGTGATGTGAAACACTGCCAGAACTAAGCGCGGCCGGCTCAGACGAGCTGC 120
 Db 79555 CAAATATGTGATGTGAAACCTGCCAGAACTAAGCGCGGCCGGCTCAGACGAGCTGC 79496
 Qy 121 CTCAGATGTAAGTGTAAACAGAGGCCAGAGGGAGGTGGTGGGGAACAATGGGCTG 180
 Db 79495 CTCAGATGTGAAGTGTAAACAGAGGGCCAGAGGGAGGTGGTGGGGAACAATGGGCTG 79436
 Qy 181 TGAGGCGCTGGGTGGCCCGGTTCCCACTCCCCCGCAGCGCCGCTCCACAGTGGTCCG 240
 Db 79435 TGAGGCGCTGGGTGGCCCGGTTCCCACTCCCCCGCAGCGCCGCTCCACAGTGGTCCG 79376
 Qy 241 CTCGGTGGTGTACGTGCGCAATTCGGGTTCAGACCCCAAGGCTGGTGTCTCCACC 300

	repeat_region	18250..18317 /rpt_family=Tc2"	
	repeat_region	18318..18624 /rpt_family="Alu"	
	repeat_region	18625..18682 /rpt_family=Tc2"	
	repeat_region	19981..20275 /rpt_family="Alu"	
	repeat_region	20755..21158 /rpt_family=ERV1"	
	repeat_region	24151..24238 /rpt_family="MIR"	
	repeat_region	25086..25464 /rpt_family=L2"	
	repeat_region	27302..27400 /rpt_family="MIR"	
	repeat_region	28301..28509 /rpt_family="MIR"	
	repeat_region	28735..29749 /rpt_family=L1"	
	repeat_region	30312..30547 /rpt_family="MaLR"	
	repeat_region	30697..30990 /rpt_family="Alu"	
	repeat_region	31833..31907 /rpt_family=MERL_type"	
	repeat_region	32075..32197 /rpt_family=L2"	
	repeat_region	32305..32561 /rpt_family="Alu"	
	repeat_region	34217..34583 /rpt_family="MaLR"	
	Query Match	88.6%; Score 2804.2; DB 9; Length 169723;	
	Best Local Similarity	94.2%; Pred. No. 0;	
	Matches 2981; Conservative	0; Mismatches 163; Indels 19; Gaps 6;	
Qy	1	AGGTTTAATCCATGAAGAAGACAGACAATTTTAAAGTGCTATTCAACAAAAATAAAGCTT	60
Dd	57205	AGGTTTAATTCATGAAGAAGACAGACAATTTTAAAGTGCTATTCAACAAAAATAAAGCTT	57146
Qy	61	CAAATATGTGATGTGAAAATGCCAGAACCTAAGCGGGCGCGGCTCAGACCAGCGCTGC	120
Dd	57145	CAAATATGTGATGTGAAAATCTGCCAGNACTAAGCGGGCGCGGCTCAGACCAGCGCTGC	57086
Qy	121	CTCAGATATAAGTGTATACAGAGGGCCAGGGAGGTGGTGGGGACACATGGGCGTG	180
Dd	57085	CTCAGATGTGAAGTGTATACAGAGGGCCAGGGAGGTGGTGGGGACACATGGGCGTG	57026
Qy	181	TGAGGCGCTGGGTGCCCGGCTCCCCAGCTCCCCCGCACCGCGCTCCACAGTGGTCCG	240
Dd	57025	TGAGGCGCTGGGTGCCCGGCTCCCCAGCTCCCCCGCACCGCGCTCCACAGTGGTCCG	56966
Qy	241	CTCCGCTTGGTGTACGTGCGCATTCGGHTCCAGACCCAGCGTGGTCTCCACC	300
Dd	56965	CTCCGCTTGGTGTACGTGCGCATTCGGHTCCAGACCCAGCGTGGTCTCCACC	56906
Qy	301	GCTTGTGTGGCCAGTGTACTGCGGTGACCGCCAGACAGCCTCGACGCTATGGAGGAG	360
Dd	56905	GCTTGTGTGGCCAGTGTACTGCGGTGACCGCCAGACAGCCTCGACGCTATGGAGGAG	56846
Qy	361	CTGTGTGCTACCCCTCAGCCCTACCTGGGCTGGTCTCGAGAGGACTACGAGAGTTCGT	420
Dd	56845	CCTGGTGTGCTACCCCTCAGCCCTACCTGGGCTGGTCTCGAGAGGACTACGAGAGTTCGT	56786
Qy	421	GCAGCACTACCTGAGAGTATGAGACCAGATGAGAATCCTTATGGTTTTCCATCGGAATG	480
Dd	56785	GCAGCACTACCTGAGAGTATGAGACCAGATGAGAATCCTTATGGTTTTCCATCGGAATG	56726
Qy	481	GTGGTATGTGACGTGCTATCGATTTTTTGTGTTCTCCTTTTTTTGTGGAGAGATTTT	540
Dd	56725	GTGGTATGTGACGTGCTATCGATTTTTTGTGTTCTCCTTTTTTTGTGGAGAGATTTT	56666

QY 541 AGATCGGTTAGGAGTGGCTTTACGTTGGAGAGAGCAAAACTTGGTGCACGCTTCT 600
Db 5665 AGATCGGTTAGGAGTGGCTTTACGTTGGAGAGAGCAAAACTTGGTGCACGCTTCT 5666
QY 601 GGAATAATTGAAGAAAATGTAACTACTTGAATAATTTAGCCTTATTTCAAAAAGAGTAT 660
Db 56605 GGAATAATTGAAGAAAATGTAACTACTTGAATAATTTAGCCTTATTTCAAAAAGAGTAT 56546
QY 661 GRAGGCTATGAGTACAGTATCTTTTGAAGGATGCCAGCTTTGAGAGGAGGCGACAGAA 720
Db 56545 GAAGGCTATGAGTACAGTATCTTTTGAAGGATGCCAGCTTTGAGAGGAGGCGACAGAA 56486
QY 721 GAAGCAGCAAGTTTGGAGGCAACTCTGTGAAGAGCTGAACAGGTCCTCAATTTCTGAATTTGAG 780
Db 56485 GAAGCAGCAAGTTTGGAGGCAACTCTGTGAAGAGCTGAACAGGTCCTCAATTTCTGAATTTGAG 56426
QY 781 GATGAATCTCTGTCTAGAAAAGACTTTAAACACAGAGAAATCTAAACATTTCTCAACAA 840
Db 56425 GATGAATCTCTGTCTAGAAAAGACTTTAAACACAGAGAAATCTAAACATTTCTCAACAA 56366
QY 841 GATGAATCTCTGTCTAGAAAAGACTTTAAACACAGAGAAATCTAAACATTTCTCAACAA 900
Db 56365 GATGAATCTCTGTCTAGAAAAGACTTTAAACACAGAGAAATCTAAACATTTCTCAACAA 56306
QY 901 CTTCAATCAAAATAGCTGAGGCAAAATCTCTGCAAGACATTTAAATGATGAGAA 960
Db 56305 CTTCAATCAAAATAGCTGAGGCAAAATCTCTGCAAGACATTTAAATGATGAGAA 56246
QY 961 CGACGGCTATACCAATTAAGAGTCTTTGAATGAAAATTTCTCAACTTCAGACAGGCCAT 1020
Db 56245 CGACGGCTATACCAATTAAGAGTCTTTGAATGAAAATTTCTCAACTTCAGACAGGCCAT 56186
QY 1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGAAAAGGAGAGTGAAGTGAATTAATAA 1080
Db 56185 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGAAAAGGAGAGTGAAGTGAATTAATAA 56126
QY 1081 CAGAAAATTAACATTTGAAGACTCCAAAGTACAGCAGAAACAGTCTCGAATGATTAAGAA 1140
Db 56125 CAGAAAATTAACATTTGAAGACTCCAAAGTACAGCAGAAACAGTCTCGAATGATTAAGAA 56066
QY 1141 AATCAATCAAGACCTTGATGACACTTGCCTCAATGATGAAGATCAGGCTGCTGCTT 1200
Db 56065 AATCAATCAAGACCTTGATGACACTTGCCTCAATGATGAAGATCAGGCTGCTGCTT 56006
QY 1201 GRAGAGACACACGGATGATGATACCTCGAATTTAGAGTGAACAGTGAATCGGAAAT 1260
Db 56005 GRAGAGACACACGGATGATGATACCTCGAATTTAGAGTGAACAGTGAATCGGAAAT 55946
QY 1261 GGTGCTTACTTAGATAATCTCCAAAGGAGCTTTGAAGAACTGATTCATGCTGCTAAG 1320
Db 55945 GGTGCTTACTTAGATAATCTCCAAAGGAGCTTTGAAGAACTGATTCATGCTGCTAAG 55886
QY 1321 TTAATGCTTTCTTAATAACCTTAGAGGAGAGAAACCAATTTATATTCAGTTGCT 1380
Db 55885 TTAATGCTTTCTTTAAAGCTTAGAGGAGAGAAACCAATTTATATTCAGTTGCT 55826
QY 1381 GAAGTTGATAAACAAGGAGAGCTTACAGAGCATATTAATAATCTTCAGACTCAACAA 1440
Db 55825 GAAGTTGATAAACAAGGAGAGCTTACAGAGCATATTAATAATCTTCAGACTCAACAA 55766
QY 1441 GCATCTTTGCAGTACAGAAACACACATTTTGAATAAGAGATCAGAGCTTCAACAGAAA 1500
Db 55765 GAATCTTTTGAATCAGAAACACATATTTTGAAGTGAAGATCAGAGCTTCAACAGAAA 55706
QY 1501 CTTAAAGTAAATGACTGAATATATCAAGAAAATGAAATGAACTCCACAGGAAATTAACA 1560
Db 55705 CTTAAAGTAAATGACTGAATATATCAAGAAAATGAAATGAACTCCACAGGAAATTAACA 55646
QY 1561 GTAGAGAAAATTTTCGGTTAGAGAAAGAGAGAAACTTTCTAAAGTAGATGAAGATC 1620
Db 55645 GTAGAGAAAATTTTCGGTTAGAGAAAGAGAGAAAGAGCTTTCTAAAGTGAAGAAAGCTC 55586
QY 1621 AGCCATGCCACTGAAGAGCTGGAGACCTATAGAAAAGCGCAAGATCTTGAAGAGAA 1680

Db 55585 AGCCGTCCTCACTGAACAGCTGGAGACCTATAGAAAAGTACCAAGATCTTGAAGAGAA 55526
QY 1681 TTGAGAGAACTATTCATCTTATCAAGGGCAGATTTATTTCCCATGAGAAAAAGCAGCAT 1740
Db 55525 TTGAGAGAACTATTCATCTTATCAAAAGCAGGTTATTTCTTACGAGAAAAGAGCAGCAT 55466
QY 1741 GATAATTTGGTGGCAGCTCGAATGCTGAAAAGAAACCTCAATGATTTTAAAGAAAAGAAAAAT 1800
Db 55465 GATAATTTGGTGGCAGCTCGAATGCTGAAAAGAAACCTCAATGATTTTAAAGAAAAGAAAAAT 55406
QY 1801 GCTCACAACAGACAAAATTAATCTGAAAACAGAGCTTAAATTTTGAACCTTTTGAAGAAAAGAT 1860
Db 55405 GCTCACAACAGACAAAATTAATCTGAAAACAGAGCTTAAATTTTGAACCTTTTGAAGAAAAGAT 55346
QY 1861 CTTATGACACTCGATGTTTCCAAATACAGCATTTTGGCAGAGAGCATTTCCCATATGCTGCC 1920
Db 55345 CTTATGACACTCGATGTTTCCAAATACAGCATTTTGGCAGAGAGCATTTCCCATATGCTGCC 55286
QY 1921 TCACATTTGGTGGCTTCAATCTGAAAACAGAGCTTTTCTCTCTCTCAACCTTTGTTG 1980
Db 55285 TCACATTTGGTGGCTTCAATCTGAAAACAGAGCTTTTCTCTCTCTCAACCTTTGTTG 55226
QY 1981 GAGGTCCTCACTCAGACTCTCACCTTTCTCCAGGGGAGAGAGAGGCTCACGAGGC 2040
Db 55225 GAGGTCCTCACTCAGACTCTCACCTTTCTCCAGGGGAGAGAGAGGCTCACGAGGC 55166
QY 2041 CCAGGGAATCTCTGGACCATCAGATTAACCAATGAAAAGAGAGAGATCAAGCTGTAAGG 2100
Db 55165 CCAGGGAATCTCTGGACCATCAGATTAACCAATGAAAAGAGAGAGATCAAGCTGTAAGG 55106
QY 2101 TTAACCGATCTCATAGGCTCCCTCTGACACTGGGCTCTGTCACTCCATCCATGGAGCAG 2160
Db 55105 TTAACCGATCTCATAGGCTCCCTCTGACACTGGGCTCTGTCACTCCATCCATGGAGCAG 55046
QY 2161 GACCGTAGGATGATTTCTCTCCGACAGGCAATCATATCTGATTCAGCCCTTCTCTCCA 2220
Db 55045 GACCGTAGGATGATTTCTCTCCGACAGGCAATCATATCTGATTCAGCCCTTCTCTCCA 54986
QY 2221 CAAAGGCAAGACAGATTTGTTCTTAATCTGGTAGACTGTGTGACAGCAGAGAGACTCAGA 2280
Db 54985 CAAAGGCAAGACAGATTTGTTCTTAATCTGGTAGACTGTGTGACAGCAGAGAGACTCAGA 54926
QY 2281 AGTTTAAATATGCTCTTTTGGATAAAATGATGGGTCAATGCCTTCAGAAAATGGAATCC 2340
Db 54925 AGTTTAAATATGCTCTTTTGGATAAAATGATGGGTCAATGCCTTCAGAAAATGGAATCC 54866
QY 2341 AGTAGAAATGATACCAAGAGATGATCTGGTAATTTAAATGTGCTGATTCATCTCTCCT 2400
Db 54865 AGTAGAAATGATACCAAGAGATGATCTGGTAATTTAAATGTGCTGATTCATCTCTCCT 54806
QY 2401 GCTGAAAATGAAGCCACTGCGCTTGTCTTCTCCACCTTGTCTTGTCTCAATCAGAGGT 2460
Db 54805 GCTGAAAATGAAGCCACTGCGCTTGTCTTCTCCACCTTGTCTTGTCTCAATCAGAGGT 54746
QY 2461 CAATTTGTTTTCAGTGGATGCAAGAGGCCCATCTTGTGAGAGAGGACCTCTTTTCCCCCA 2520
Db 54745 CAATTTGTTTTCAGTGGATGCAAGAGGCCCATCTTGTGAGAGAGGACCTCTTTTCCCCCA 54686
QY 2521 CCTCTCTCAGAGAGCCATGTTTGGAGCTTCTCGAGATTTATTTTCCACCAAGGGATTTCCCA 2580
Db 54685 CCTCTCTCAGAGAGCCATGTTTGGAGCTTCTCGAGATTTATTTTCCACCAAGGGATTTCCCA 54626
QY 2581 GGTCCACCACTGCTTCCATTTGCAATGAGAAATGTCTATTCACCGAGGGGTTTCTCTCT 2640
Db 54625 GGTCCACCACTGCTTCCATTTGCAATGAGAAATGTCTATTCACCGAGGGGTTTCTCTCT 54566
QY 2641 TACCTTCCCCCAAGACCTGGATTTTCCCCCCCCACCATTTCTGAAGGTAGAGTGAAG 2700
Db 54565 TACCTTCCCCCAAGACCTGGATTTTCCCCCCCCACCATTTCTGAAGGTAGAGTGAAG 54508
QY 2701 TTTCCCTCAGGTTTGAATTCACCTTTCAATGAGGCTGTCTACTGAAACATTCAGAACCA 2760

Db 78579 GGACTAATTTGAAGAAAATGTAACCTACTTTGAAAAGTTTAGCCTTATTTCAAAAAGAGTAT 78638
Qy 661 GAAGGCTATGAAGTAGAGTCACTCTTTAGAGGATGCGCAGCTTTGAGAAGGAGCGACAGAA 720
Db 78639 GAAGGCTATGAAGTAGAGTCACTCTTTAGAGGATGCGCAGCTTTGAGAAGGAGCGCAGAGAA 78698
Qy 721 GAAGCAGGAAGTTTGAGGCAACCTCTGAAAAGCTGAAAGAGTCCAAATTCCTGAAGCTGAG 780
Db 78699 GAAGCAGGAAGTTTGAGGCAACCTCTGAAAAGCTGAAAGAGTCCAAATTCCTGAAGCTGAG 78758
Qy 781 GATGAATCTCTGCTAGAAAAGAACTTAAACAAAGAGAAAATCTAAACATTTCTCAACAA 840
Db 78759 GATGAATCTCTGCTAGAAAAGAACTTAAAGAGAGAAAATCTAAACATTTCTCAACAA 78818
Qy 841 GATGAATGATGGCGGATATTTCAAAAAGTATACAGTCTCTAGAGAGTGAAGTCAAAATCC 900
Db 78819 GATGAATGATGGCGGATATTTCAAAAAGTATACAGTCTCTAGAGAGTGAAGTCAAAATCC 78878
Qy 901 CTCAAAATCAAAATAGCTGAAGCAAAATCATCTGCAAGACATTTTAAATGAGTGAAGAA 960
Db 78879 CTCAAAATCAAAATAGCTGAAGCAAAATCATCTGCAAGACATTTTAAATGAGTGAAGAA 78938
Qy 961 CGACGGGCTATACAAATAAAGATGCTTTGAATGAATAATCTCAACTTCAGACAGCCAT 1020
Db 78939 CGACGGGCTATACAAATAAAGATGCTTTGAATGAATAATCTCAACTTCAGACAGCCAT 78998
Qy 1021 AAACAGCTTTTCAGCAAGAGCTGAAGTATGAAAAGGAGAGTGAAGTGAATTAATAA 1080
Db 78999 AAACAGCTTTTCAGCAAGAGCTGAAGTATGAAAAGGAGAGTGAAGTGAATTAATAA 79058
Qy 1081 CAGAAAATTAACATTTGAAGACTCAAAGTACACGAGAACAAAGTTCTGAATGAATAAGAA 1140
Db 79059 CAGAAAATTAACATTTGAAGACTCAAAGTACACGAGAACAAAGTTCTGAATGAATAAGAA 79118
Qy 1141 AATCAGATCAAGACCTGACTGACACTTCCCAATGATGAAGATCAGGCTGCTGCTT 1200
Db 79119 AATCAGATCAAGACCTGACTGACACTTCCCAATGATGAAGATCAGGCTGCTGCTT 79178
Qy 1201 GAAGAAGACAAACGGATGATGATACTCGAAATAGAAAGTGAACAGTGAATCGGAAAT 1260
Db 79179 GAAGAAGACAAACGGATGATGATACTCGAAATAGAAAGTGAACAGTGAATCGGAAAT 79238
Qy 1261 GGTGCTTACTAGATATCTCCAAAGGAGCTTTGAAGAACTGATTCATGCTGCTAAG 1320
Db 79239 GGTGCTTACTAGATATCTCCAAAGGAGCTTTGAAGAACTGATTCATGCTGCTAAG 79298
Qy 1321 TTAATGCTCTTTTAAAGACTTAGAAGGAGAAAGAAACCAATTTATATTCAGTTGCT 1380
Db 79299 TTAATGCTCTTTTAAAGACTTAGAAGGAGAAAGAAACCAATTTATATTCAGTTATCT 79358
Qy 1381 GAAGTTGATTAACAAAGGAGAGCTTACAGAGCATATTAATAATCTTCAGACTGAACAA 1440
Db 79359 GAAGTTGATTAACAAAGGAGAGCTTACAGAGCATATTAATAATCTTCAGACTGAACAA 79418
Qy 1441 GCATCTTTGAGTCAGAAAACACACATTTTGAATAATGAGAAATCAGAGCTTCAACAGAAA 1500
Db 79419 GCATCTTTGAGTCAGAAAACACATATTTTGAAGTGAAGTCAAGAGCTTCAACAGAAA 79478
Qy 1501 CTTAAAGTATGATGATATTAATCAAGAAAATGAATGAATACTCCACAGAAAATTAACA 1560
Db 79479 CTTAAAGTATGATGATATTAATCAAGAAAATGAATGAATACTCCACAGAAAATTAACA 79538
Qy 1561 GTAGAGAAAATTAATCGGTTAGAAAAGAGAGAAAATTTCTAAAGTAGATGAAGAATC 1620
Db 79539 GTAGAGAAAATTAATCGGTTAGAAAAGAGAGAAAATTTCTAAAGTAGATGAAGAATC 79598
Qy 1621 AGCCATGCCATGAAGAGCTGGAGACTATAGAAAAGCGGACCAAGATCTTTGAGAGAGAA 1680
Db 79599 AGCCATGCCATGAAGAGCTGGAGACTATAGAAAAGCGGACCAAGATCTTTGAGAGAGAA 79658
Qy 1681 TTGAGAGAACTATTCATCTTATCAAGGAGAGATTTTCCCATGAGAAAAGAGACAT 1740
Db 79659 TTGAGAGAACTATTCATCTTATCAAGGAGAGATTTTCCCATGAGAAAAGAGACAT 79718

Qy 1741 GATAATTTGGTGGAGCTCGGAATGCTGAAAAGAAACCTCAATGATTTAGGAAAAGAAAT 1800
Db 79719 GATAATTTGGTGGAGCTCGGAATGCTGAAAAGAAACCTCAATGATTTAGGAAAAGAAAT 79778
Qy 1801 GCTCAACAACAGACAAAATTAACCTGAAAACAGAGCTTAAATTTGAACTTTTGAAGAAAAGAT 1860
Db 79779 GCTCAACAACAGACAAAATTAACCTGAAAACAGAGCTTAAATTTGAACTTTTGAAGAAAAGAT 79838
Qy 1861 CTTATGCACTCGATGTTCCAAATACAGATTTTGGCAGAGAGCATTTCCCATATGTTCC 1920
Db 79839 CTTATGCACTCGATGTTCCAAATACAGATTTTGGCAGAGAGCATTTCCCATATGTTCC 79898
Qy 1921 TCACCATTTGGTGGCCTTCATCTGAAAACAGAGCTTTTCTCTCTCTCTCAACTTTGTTG 1980
Db 79899 GCACCATTTGGTGGCCTTCATCTGAAAACAGAGCTTTTCTCTCTCTCTCAACTTTGTTG 79958
Qy 1981 GAGGTCACATCAGACTCTCACCTTTCCAGGGGAGAGAGAGAGCTCAGAGGC 2040
Db 79959 GAGGTCACATCAGACTCTCACCTTTCCAGGGGAGAGAGAGAGCTCAGAGGC 80018
Qy 2041 CCAGGATCTCTGGAACCATCAGATTTACCAATGAAAAGAGAGAAATCAAGCTGTGATAG 2100
Db 80019 CCAGGATCTCTGGAACCATCAGATTTACCAATGAAAAGAGAGAAATCAAGCTGTGATAG 80078
Qy 2101 TTAACCGATCTCATAGGGTCTCTCTGACACTGGGTCTCTGTCACCTCCATGGGACAG 2160
Db 80079 TTAACCGATCTCATAGGGTCTCTCTGACACTGGGTCTCTGTCACCTCCATGGGACAG 80138
Qy 2161 GACCGTAGGATGATTTCTCCGACAGGCAATCATATCTGATTCAGGCTTTCTCTCCA 2220
Db 80139 GACCGTAGGATGATTTCTCCGACAGGCAATCATATCTGATTCAGGCTTTCTCTCCA 80198
Qy 2221 CAAAGGCAAGACAGATTTTGTCTTAATTTCTGGTAGACTGTCTGACGACGAGAACTCAGA 2280
Db 80199 CAAAGGCAAGACAGATTTTGTCTTAATTTCTGGTAGACTGTCTGACGACGAGAACTCAGA 80258
Qy 2281 AGTTTAAATGACCTTTCTGGATTAATAATGATGGGTCAATGCTTCAGAAATGGAATCC 2340
Db 80259 AGTTTAAATGACCTTTCTGGATTAATAATGATGGGTCAATGCTTCAGAAATGGAATCC 80318
Qy 2341 AGTAAATGATACCAAGATGATCTTGGTAAATTTAAATGTGCTGATTCATCTCTCCCT 2400
Db 80319 AGTAAATGATACCAAGATGATCTTGGTAAATTTAAATGTGCTGATTCATCTCTCCCT 80378
Qy 2401 GCTGAAAATGAAGCACTGGCCCTGGCTTTGTTCTCCACCTCTTGTCTCAATCAGAGT 2460
Db 80379 GCTGAAAATGAAGCACTGGCCCTGGCTTTGTTCTCCACCTCTTGTCTCAATCAGAGT 80438
Qy 2461 CCATTTGTTTCCAGTGAATGCAAGAGGCCATTTCTGAGAAGAGGACCTCTTTCCCCCA 2520
Db 80439 CCATTTGTTTCCAGTGAATGCAAGAGGCCATTTCTGAGAAGAGGACCTCTTTCCCCCA 80498
Qy 2521 CTTCTCTCAGGAGCCATGTTTGGAGCTTCTCGAGTTATTTTCCACCAAGGGATTTCCCA 2580
Db 80499 CTTCTCTCAGGAGCCATGTTTGGAGCTTCTCGAGTTATTTTCCACCAAGGGATTTCCCA 80558
Qy 2581 GGTCCACCACTGCTTCCATTTGCAATGAGAAATGTCATCCACGAGGGGTTTCTCTCT 2640
Db 80559 GGTCCACCACTGCTTCCATTTGCAATGAGAAATGTCATCCACGAGGGGTTTCTCTCT 80618
Qy 2641 TACCTTCCCCCAAGACCTGGAATTTTCCCCCAAGGGGTTTCTGAAAGTAGAGTGA 2700
Db 80619 TACCTTCCCCCAAGACCTGGAATTTTCCCCCAAGGGGTTTCTGAAAGTAGAGTGA 80676
Qy 2701 TTCCCTCTCAGGTTTGAATTCACCTTCAATGAGCTGCTACTGAAATCAGAGAGAGAG 2760
Db 80677 TTCCCTCTCAGGTTTGAATTCACCTTCAATGAGCTGCTACTGAAATCAGAGAGAGAG 80728
Qy 2761 CAAGAAACCTGCAATATTTTGTCTCTCTTCAAAAGTAAATTTTGTGATCTCATTTTCA 2820
Db 80729 CAAGAAACCTGCAATATTTTGTCTCTCTTCAAAAGTAAATTTTGTGATCTCATTTTCA 80788

```

QY 2821 GTTTAAGTAACGCTGTTACTTTAAGTGAATACATTTTGGCTCAAAATGAAGCTTAATGGA 2880
Db 80789 GTTTAAGTAACGCTGTTACTTTAAGTGAATACATTTTGGCTCAAAATGAAGCTTAATGGA 80848
QY 2881 ATTATAATTTCTCAGGATAGTATTTTGTAAATAAAGATGATTTAAATATGAATCTTATGAG 2940
Db 80849 ATTATAATTTCTCAGGATAGTATTTTGTAAATAAAGATGATTTTAAATAGGAATCTTATGAG 80908
QY 2941 TAAATATTTCGAATTTTATTTTCTAGATCAATATACTATTTTAAATTTGGTGAATTAATCC 2994
Db 80909 TAAATCAATCCATTTTATTTTCTAGATCAATATACTATTTTAAATTTGGTGAATTAATCC 80968
QY 2995 ACTATTATATAACATAGTGGAGTTTATATATATGTAATCTTCCAGGTGGGGAGGCTTT 3054
Db 80969 ACTGTTATAGAACAATATAGGAGTTTATATATATGTAATCTTCCAGGTGGGGAGGCTTT 81028
QY 3055 AAATCTCTGAAGTCT---GTGCTCTTATGCGCAAGAACTGATTTACTGTGTTGTGGACAA 3111
Db 81029 AAATCTCTGAAGTCTGTTGGTGTCTTCATGCGCAAGAACTGATTTACTGTGTTGTAGATAA 81088
QY 3112 ATGCGAAGTAACCTTATGCTTAAATATAATTTATAGCTTGAATTA 3154
Db 81089 ATGCGAAGTAACCTTATGCTTAAATATAATTTATAGCTTGAATTA 81131

RESULT 8
AC073264
LOCUS Homo sapiens BAC clone RP11-298A10 from 7, complete sequence.
DEFINITION AC073264
ACCESSION AC073264.5 GI:13443255
VERSION HTG.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE
AUTHORS Sulston,J.E. and Wilson,R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE
AUTHORS Abbott,A., Doebber,A., Martinka,S. and Tomlins,C.
TITLE The sequence of Homo sapiens BAC clone RP11-298A10
JOURNAL Unpublished (2001)
REFERENCE
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-2000) Genome Sequencing Center, Washington University School of Medicine, 444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (24-MAR-2001) Genome Sequencing Center, Washington University School of Medicine, 444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (09-MAY-2001) Department of Genetics, Washington University, 444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE
AUTHORS Wilson,R.
TITLE Direct Submission
JOURNAL Submitted (15-OCT-2003) Department of Genetics, Washington University, 444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT
On Mar 24, 2001 this sequence version replaced gi:11192200.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu

```

Contact: sapiens@watson.wustl.edu
 ----- Summary Statistics
 Center project name: H_NH0298A10

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred score >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GR7/CHR7>, send mailto:egreen@nhgri.nih.gov, or see <http://genome.wustl.edu>

SOURCE INFORMATION:
 The RP11-11 human BAC library was made from the blood of one male donor, as described by Osogawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries for construction of 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
 VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the right is AC024155. Actual start of this clone is at base position 1 of RP11-298A10 actual end is at base position 200123 of RP11-298A10.

FEATURES	Location/Qualifiers
source	1..200123
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="7"
	/map="7"
repeat_region	/Clones="RP11-298A10"
	/Clone_lib="RPC1-11"
	259..763
	/rpt_family="L1"
repeat_region	779..1347
	/rpt_family="ERV1"
repeat_region	1345..2101
	/rpt_family="ERV1"
repeat_region	2321..2728
	/rpt_family="ERV1"
repeat_region	2745..2945
	/rpt_family="L1"
repeat_region	3043..3082
	/rpt_family="A-rich"
repeat_region	3102..3293
	/rpt_family="L2"
repeat_region	3340..3452
repeat_region	3478..4173
	/rpt_family="L2"
repeat_region	5403..5702
	/rpt_family="Alu"
repeat_region	6528..6566
	/rpt_family="L2"
repeat_region	Product="rRNA-28S_rRNA-related"
repeat_region	3478..4173
repeat_region	5403..5702
repeat_region	6528..6566
repeat_region	Product="rRNA-28S_rRNA-related"

```

repeat_region 7084. .7363
/rpt_family="Alu"
repeat_region 8128. .8452
/rpt_family="MIR"
repeat_region 8613. .9033
/rpt_family="MaLR"
repeat_region 9081. .9193
/rpt_family="L1"
repeat_region 10668. .10096
/rpt_family="(CA)n"
repeat_region 10121. .10721
/rpt_family="ERV1"
repeat_region 10788. .10943
/rpt_family="MIR"
repeat_region 11143. .11902
/rpt_family="ERV1"
repeat_region 11348. .12146
/rpt_family="MaLR"
repeat_region 12147. .12832
/rpt_family="L1"
repeat_region 12833. .13301
/rpt_family="MaLR"
repeat_region 13390. .13824
/rpt_family="MaLR"
repeat_region 14968. .15129
/rpt_family="MIR"
repeat_region 15148. .15245
/rpt_family="L1"
repeat_region 15277. .15433
/rpt_family="L1"
repeat_region 15434. .15745
/rpt_family="Alu"
repeat_region 15746. .16513
/rpt_family="L1"
repeat_region 16521. .16612
/rpt_family="MIR"
repeat_region 17075. .17710
/rpt_family="L2"
repeat_region 17711. .18004
/rpt_family="Alu"
repeat_region 18005. .18067
/rpt_family="L2"
repeat_region 18068. .18390
/rpt_family="Alu"
repeat_region 18391. .18442
/rpt_family="L2"
repeat_region 18929. .19301
/rpt_family="MaLR"
repeat_region 19952. .20213
/rpt_family="L2"
repeat_region 20215. .20243
/rpt_family="(TGA)n"
repeat_region 20312. .20346
/rpt_family="AT-rich"
repeat_region 20566. .21492
/rpt_family="L1"
repeat_region 21493. .22063
/rpt_family="ERVX"
repeat_region 22064. .24908
/rpt_family="L1"
repeat_region 24909. .25219
/rpt_family="Alu"
repeat_region 25220. .26128
/rpt_family="L1"
repeat_region 26226. .26550
/rpt_family="MaLR"
repeat_region 28925. .29206
/rpt_family="Alu"
repeat_region 29821. .30421
/rpt_family="L1"
repeat_region 30504. .30842
/rpt_family="L1"
repeat_region 30846. .30930

```

```

/rpt_family="MaLR"
repeat_region 30936. .31014
/rpt_family="MIR"
repeat_region 32007. .32232
/rpt_family="L1"
repeat_region 32232. .32696
/rpt_family="L1"
repeat_region 32697. .32995
Query Match 88.6%; Score 2804.2; DB 9; Length 200123;
Best Local Similarity 94.2%; Pred. No. 0;
Matches 2981; Conservative 0; Mismatches 163; Indels 19; Gaps 6;
QY 1 AGGTTTAAATCCATGAAGAAGACAGCAATTTTAAAGTGTATTACCAAAAAATAAGCTT 60
Db 120786 AGGTTTAAATCCATGAAGAAGACAGCAATTTTAAAGTGTATTACCAAAAAATAAGCTT 120845
QY 61 CAAATAATGTATGTGAAAACTGCCAGAACTAAAGCGCGCGCGCTCAGACCAAGCGCTGC 120
Db 120846 CAAATAATGTATGTGAAAACTGCCAGAACTAAAGCGCGCGCGCTCAGACCAAGCGCTGC 120905
QY 121 CTCAGGATGTAAGTGTAAACAAGAGGGCCAGGGAGGTGGTGGGGGACAAACATGGGCTG 180
Db 120906 CTCAGGATGTAAGTGTAAACAAGAGGGCCAGGGAGGTGGTGGGGGACAAACATGGGCTG 120965
QY 181 TGAGGCTGTGGGTGCCCGCTTCCAGCTCCCGCCAGCCCGCTCCACAGTGGTCCG 240
Db 120966 TGAGGCTGTGGGTGCCCGCTTCCAGCTCCCGCCAGCCCGCTCCACAGTGGTCCG 121025
QY 241 CTCGGTTGGTTGCACGTGCGCATTCGGGTTCAGACCCCAAGGCTGGTCTCCACC 300
Db 121026 CTCGGTTGGTTGCACGTGCGCATTCGGGTTCAGACCCCAAGGCTGGTCTCCACC 121085
QY 301 GCTTGTTCGGCCAGTGTACTGCGGTGACCGCCAGAGCAGCCTCGAGCTATGGAGGAG 360
Db 121086 GCTTGTTCGGCCAGTGTACTGCGGTGACCGCCAGAGCAGCCTCGAGCTATGGAGGAG 121145
QY 361 COTGCTGTACCCCTCAGCCCTACCTGGGGCTGGTCTGGAGGAGCTACGAGAGTTGTG 420
Db 121146 CCGGTGCTACCCCTCAGCCCTACCTGGGGCTGGTCTGGAGGAGCTACGAGAGTTGTG 121205
QY 421 GCAGCACTACCTGAGAGTATGAGACCAAGATCAGAACTCTTATGGTTTCCATCGGAATG 480
Db 121206 GCAGCACTACCTGAGAGTATGAGACCAAGATCAGAACTCTTATGGTTTCCATCGGAATG 121265
QY 481 GTGGTATGTCAGCTGTTATTGGATTTTGTCTCTCTTTTGTGGAGAGTTT 540
Db 121266 GTGGTATGTCAGCTGTTATTGGATTTTGTCTCTCTTTTGTGGAGAGTTT 121325
QY 541 AGATCGGTAGGAGTCGGCTTTACGTGGGAAGAGAGCAAAAACCTTGTGCAACGCTTTCT 600
Db 121326 AGATCGGTAGGAGTCGGCTTTACGTGGGAAGAGAGCAAAAACCTTGTGCAACGCTTTCT 121385
QY 601 GGACTAATGAAGAAAAATGTAACCTACTTGAATAATTTAGCCTTATTCAAAAAGAGTAT 660
Db 121386 GGACTAATGAAGAAAAATGTAACCTACTTGAATAATTTAGCCTTATTCAAAAAGAGTAT 121445
QY 661 GAAGGCTATGAAGTAGAGTCACTTTTAGAGTAGCCAGCTTTGAGAGGAGGAGCAGAGAA 720
Db 121446 GAAGGCTATGAAGTAGAGTCACTTTTAGAGTAGCCAGCTTTGAGAGGAGGAGCAGAGAA 121505
QY 721 GAAGCAGAGAGTTTGGAGGCAACCTGTGAAAGCTGAACAGTCCCAATCTGAATTCAG 780
Db 121506 GAAGCAGAGAGTTTGGAGGCAACCTGTGAAAGCTGAACAGTCCCAATCTGAATTCAG 121565
QY 781 GATGAATCCTCTGTCTAGAAAAAGACTTAAACAAGAGAAATCTAAACATTTCTCAACAA 840
Db 121566 GATGAATCCTCTGTCTAGAAAAAGACTTAAACAAGAGAAATCTAAACATTTCTCAACAA 121625
QY 841 GATGAATGATCGCGGATTTTCAAAAAAGTATACAGTCTCTAGAGAGTGAAGTCAAAATCC 900
Db 121626 GATGAATGATCGCGGATTTTCAAAAAAGTATACAGTCTCTAGAGAGTGAAGTCAAAATCC 121685

```

QY	901	CTCAAAATCAAAATAGCTGAAGCCAAATATCATCTGCAAGACATTTTAAAAATGAGTGAAGAA	960
Db	121686	CTCAAAATCACAAATAGCTGAAGCCAAATATCATCTGCAAGACATTTTAAAAATGAGTGAAGAA	121745
QY	961	CGACGGCTATAGCAATAAAAGATGCTTTGAAATGAAAAATTTCTCAATCTCAGACAAAGCCAT	1020
Db	121746	CGACGGCTATAGCAATAAAAGATGCTTTGAAATGAAAAATTTCTCAATCTCAGACAAAGCCAT	121805
QY	1021	AAACAGCTTTTTTCAGCAAGAGCTGAAGTATGGAAGGAGAGTGAAGTGAATTAATAAA	1080
Db	121806	AAACAGCTTTTTTCAGCAAGAGCTGAAGTATGGAAGGAGAGTGAAGTGAATTAATAAA	121865
QY	1081	CAGAAAAATAACATTTTGAAGACTCCAAAGTACACGCGAGAACAAAGTTCTGTAATGATAAAGAA	1140
Db	121866	CAGAAAAATAACATTTTGAAGACTCCAAAGTACACGCGAGAACAAAGTTCTGTAATGATAAAGAA	121925
QY	1141	AATCACATCAAGACCGCTGACTGGACATCTCCCAATGATGAAGATCAGGCTGCTGTGCTT	1200
Db	121926	AATCACATCAAGACCGCTGACTGGACATCTCCCAATGATGAAGATCAGGCTGCTGTGCTT	121985
QY	1201	GAGAGAGACAAACGGATGATGATAAACCCTGGAATTAGAAGTGAACAGTGAATCGGAAAAAT	1260
Db	121986	GAGAGAGACAAACGGATGATGATAAACCCTGGAATTAGAAGTGAACAGTGAATCGGAAAAAT	122045
QY	1261	GGTGCTTACTTAGATAAATCCTCCAAAGAGCTTTGAAGAACTGATTGATCTGCTGTAG	1320
Db	122046	GGTGCTTACTTAGATAAATCCTCGAAAGAGCTTTGAAGAACTGATTGATCTGCTGTAG	122105
QY	1321	TTAAATGCTCTTTTAAAAACCTTAGAGGAGAAAAACCAAAATTTATATTTCAAGTTGTCT	1380
Db	122106	TTAAATGCTCTTTTAAAAACCTTAGAGGAGAAAAACCAAAATTTATATTTCAAGTTGTCT	122165
QY	1381	GAGTTTGATAAACAAGAGAGAGCTTACAGAGCATATTAAGAATCTTCAGACTCAACAA	1440
Db	122166	GAGTTTGATAAACAAGAGAGAGCTTACAGAGCATATTAAGAATCTTCAGACTCAACAA	122225
QY	1441	GCATCTTTCAGTCAAGAAACACACATTTTGAATAAGAGATCAGAGCTTCAACAGAA	1500
Db	122226	GCATCTTTCAGTCAAGAAACACATATTTTGAATAAGAGATCAGAGCTTCAACAGAA	122285
QY	1501	CTTAAAGTAAATGACTGAAATTTATCAAGAAATGAAATGAACTCCACAGGAAATTAACA	1560
Db	122286	CTTAAATAAATGACTGAAATTTATCAAGAAATGAAATGAACTTCACAGGAAATTTAACA	122345
QY	1561	GTAGAGAAAAATATCGGTTAGAGAAAGAGAGAACTTTCTTAAAGTAGATGAAAAGATC	1620
Db	122346	GTGAGAGAAAAATACCGAATAGAGAAAGAGAGAGCTTTCTAGAGTGGAGAGAAAGCTC	122405
QY	1621	AGCCATTGCCACTGAAGAGCTGGAGACCTTATAGAAGCGAGCCAAAGATCTTTGAAGAGAA	1680
Db	122406	AGCCGTGCCACTGAACAGCTGGAGACCTTATAGAAGCTAGCCAAAGATCTTGAAGAGAA	122465
QY	1681	TTGGAGAGAACTATTCATTTCTATCAAGGCGAGATATTTCCCATGAGAAAAAGCACAT	1740
Db	122466	TTGGAGAGAACTGTTCATTTTTTATCAAAAGCAGGTTATTTCCCTACAGAAAGAGAGACAT	122525
QY	1741	GATAATTGGTTGGCAGCTCGGAATGCTGAAGAAACCTCAATGATTTTAAAGGAAAGAAAT	1800
Db	122526	GATAATTGGTTGGCAGCTCGGACTGCTGAAGAAACCTCAGTGATTTTAAAGGAAAGAAAT	122585
QY	1801	GCTCACACAGACAAAAATTAATGAAAACAGAGCTTTAAATTTGAACTTTTAGAAAAAGAT	1860
Db	122586	GCTCACACACAAAAATTAATGAAAACAGAGTTGAAATTTGAACTTTTAGAAAAAGAT	122645
QY	1861	CCTTATGCACTCGATGTTCCAAATACAGCATTTGGCAGAGAGCATTCGCCATATGGTCCC	1920
Db	122646	CCTAATGCACTTGATGTTTCAAAATACAGCATTTGSCAGAGAGCATTCGCCATATGGTCCC	122705
QY	1921	TCACCATTTGGTTGGCTTCATCTGNAACAGAGCTTTTCTCTCTCTCCACTTTGTTG	1980
Db	122706	GCAACCATTTGGTTGGCTTCATCTGNAACAGAGCTTTTCTCTCTCTCCACTTTGTTG	122765
QY	1981	GAGGGTCCACTCAGACTCACTTCCTTTGCTTCCAGGGGGAGAGGAAAGGCTCACGAGGC	2040

Db	122766	GAGGATCCACTGGACCTCTCACCTGTGCTTCCGAGGGAGGAGAGGCCAAGAGGC	122825
Qy	2041	CCAGGGAATCCTCTGGACCATCAGATTACCAATGAAAGAGGAGAAATCAAGCTGTGATAGG	2100
Db	122826	CCAGGGAATCCCTCTGGACCATCAGATTACCAATGAAAGAGGAGAAATCAAGCTGTGATAGG	122885
Qy	2101	TTAACCGATCCTCATAGGGCTCCTCTGACACCTGGGTCTCTGTCACTCCTCATGGGACCAAG	2160
Db	122886	TTAACCGATCCTCATAGGGCTCCTCTGACACCTGGGTCTCTGTCACTCCTCATGGGACCAAG	122945
Qy	2161	GACCGTAGGATGATTTTCCTCCGACAGGACAATCATATCCTGATCAGGCCCTTCTCTCCA	2220
Db	122946	GACTGTAAAGATGATTTTCCTCCGACAGGACAATCATATCCTGATCAGGCCCTTCTCTCCA	123005
Qy	2221	CAAAGGCAAGACAGATTTTGTCTTAATCTCTGGTAGACTGTCTGACAGACAGAACTCAGA	2280
Db	123006	CAAAGGCAAGACAGATTTTGTCTTAATCTCTGGTAGACTGTCTGACAGACAGAACTCAGA	123065
Qy	2281	AGTTTTTAATATGCTCTCTTTGGATAAAATGGATGGGTCAATGCTTCAGAAATCGAATCC	2340
Db	123066	AGTTTTTAATATGCTCTCTTTGGATAAAATGGATGGGTCAATGCTTCAGAAATCGAATCC	123125
Qy	2341	AGTAGAATGATACCAAGATGATCTTTGGTAATTTAAATGTGCTGATTCATCTCTCCTCT	2400
Db	123126	AGTAGAATGATACCAAGATGATCTTTGGTAATTTAAATGTGCTGATTCATCTCTCCTCT	123185
Qy	2401	GCTGAAATGAAGCCATCGGCCCTGGCTTTGCTCTCCACTCTTGCTCCATCAGAGGT	2460
Db	123186	GCTGAAATGAAGCCATCGGCCCTGGCTTTGCTCTCCACTCTTGCTCCATCAGAGGT	123245
Qy	2461	CCATTGTTCCAGTGGATGCAAGAGGCCCATTTCTGAGAGAGACCTCTTTCCGCCCA	2520
Db	123246	CCATTGTTCCAGTGGATGCAAGAGGCCCATTTCTGAGAGAGACCTCTTTCCGCCCA	123305
Qy	2521	CCTCCTCCAGAGCCATGTTTGGAGCTTCTCGAGATTATTTTCCACCAAGAGGATTTCCCA	2580
Db	123306	CCTCCTCCAGAGCCATGTTTGGAGCTTCTCGAGATTATTTTCCACCAAGAGGATTTCCCA	123365
Qy	2581	GGTCCACCATCTGCTCCATTGTCCATTTGCAATGAGAAATCTCTATCCACGAGGGTCTTCTCCT	2640
Db	123366	GGTCCACCATCTGCTCCATTGTCCATTTGCAATGAGAAATCTCTATCCACGAGGGTCTTCTCCT	123425
Qy	2641	TACCTTCCCCAAGACCTGGATTTTTCCCCCACCACCCACATTTCTGAAGTAGAAGTCAG	2700
Db	123426	TACCTTCCCCAAGACCTGGATTTTTCCCCCACCACCCACATTTCTGAAGTAGAAGTCAG	123483
Qy	2701	TTCCCTCAGGTGTTGATTCACCTTCAAATGAGCTGCTACTGAACATCCAGAACCAAG	2760
Db	123484	TTCCCTCAGGTGTTGATTCAGCTTTAA-----TGCTACTGAACATCCAGAACCA--	123535
Qy	2761	CAAGAACTCAGCAATATTTTGTCTCTTCAAAAGTAATTTTGACTGATCTCATTTTCA	2820
Db	123536	CAAGAACTCAGCAATATTTTGTCTCTTCAAAAGTAATTTTGACTGATCTCATTTTCA	123595
Qy	2821	GTTTAAGTAACTGCTGTGTACTTAAGTGATTAACATTTTGTCTCAAAATGAAGCTTAAATGA	2880
Db	123596	GTTTAAGTAACTGCTGTGTACTTAAGTGATTAACATTTTGTCTCAAAATGAAGCTTAAATGA	123655
Qy	2881	ATTATAATCTCAGATAGTATTTTGTAAATAAAGATGATTTAAATATGAATCTTATGAG	2940
Db	123656	ATTATAATCTCAGATAGTATTTTGTAAATAAAGATGATTTAAATATGAATCTTATGAG	123715
Qy	2941	TAAATATTTCAATTT--TTATTTAGCGGTATAACTATTTCAATTTG-----ATTAAATCC	2994
Db	123716	TAAATCATTTCAATTTTATTTCTAGATCATATACTATTTTAAATTTGGTGAATTTAATCC	123775
Qy	2995	ACTATTATATAACAATAGTGGAGTTTTTATATATGTAATCTTTTCCAGTGGGAGGCTTT	3054
Db	123776	ACTGTATATGAACAATTAATGGAGTTTTTATATATGTAATCTTTTCCAGTGGGAGGCTTT	123835
Qy	3055	AAATTTCTGAAGTCT---GTGTCCTTTATGCCAAGAACTGTATTTTACTGTGGTGTGCAACA	3111


```

123836  AAATCTCAAAGTGTGGTGTCTTCATGCCCAAGNACTGTATTACCTGTGGTGTAGATAA 123893
      QY      3112  ATGTGAAGTAACCTTTATGCTTAATAATAATATATAGTTGATTTA 3154
      DB      123896  ATGTGAAGTAACCTTTATGCTTAATAATAATAATATATAGTTGA 123938

RESULT 9
HSU94780
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3676)
Heckel,D., Brass,N., Fischer,U., Blin,N., Steudel,I., Tureci,O.,
Fackler,O., Zang,K.D. and Meese,E.
cDNA cloning and chromosomal mapping of a predicted coiled-coil
proline-rich protein immunogenic in meningioma patients
Hum. Mol. Genet. 6 (12), 2031-2041 (1997)
97472454
9356211
2 (bases 1 to 3676)
Heckel,D., Brass,N. and Meese,E.U.
Direct Submission
Submitted (20-MAR-1997) Inst. for Human Genetics, University of
Saarland; Medical School, Oskar-Orth-Strasse, Homburg, Saarland
66421, Germany
FEATURES
Location/Qualifiers
1. 3676
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="meningioma"
/note="PCR and in-situ hybridization mapping locate the
gene to chromosomes 2,3,6,7,9,13,14"
1. 3676
/gene="MEA6"
315..2729
/gene="MEA6"
CDS
/note="meningioma-expressed antigen 6; this sequence is
similar to MEA11, GenBank Accession Number U73682, but is
longer; it may be an alternate splice variant or it may be
from a separate gene"
/codon_start=1
/product="MEA6"
/protein_id="AAB96593.1"
/db_xref="GI:2231999.1"
/translation="MEBFGATPQYLGILLLELRVVAALPEGMRRDSNLGYFPWELV
ICAAVGFVAFLLFWRSFRVRRLYVGREKLAUMLSGLEILBEKSKLLEKFSVQKE
YEGVEVSEKDAQFKEATEAQALCEATCEKLNRSNEDEILCLEKELEKSKHS
EQELMDALIKRIQSLEDESXSLKQVAEAKMTFQIPMNEERLKIATKDALNENSL
QESQOLLQAEVAVKEQVLSNOKQVTFEDSKVHAQVNDKESHTKTULRELLKMD
WAALGKLTDDNLELMNSESNGAVLDNPPKGLKLIHAAKNLSLKTLEGERN
QYIOLSVDTKTELTEHINKLOTEAQSLQSENTHPENENOKLOKLVMTLEGRN
EMKLRKLTENVLEKEELKLVDSKISHATELETYKRAKDLLEELERTHSYQ
GKIIISHEKADNDMLAANARNLNDLRKNAHRNQLTTELKFLLEKSDPTALDVF
NTAFSGHSPYGGPGLWMPSSSETAFVLPPTLEGLPTLSPLPGGSGRSGRGNPL
DHOITNREGSSCDRLTDPHRALSDTGLSPDQDRMMFPFPGOSYPSDALPPQOR
DRFCNSGRLSGPAELRSFNPNPSLDKMDGMPSESSNRNDTKDGLNLNVPDSLPA
ENEAATGPGVFPPLAVRGPPLFVDARGPFLRRGPPPPPPGAMGASRDYPPGDF
PGPPAPFAMRNVPYPPRGFPYLPFRPGFTPPPHSGRSEFTPSGLIPSPNEPATEHP
BQQQTT"
1354..1982
/misc_feature
/gene="MEA6"
/note="possible alternatively spliced exon"
3135..3140
/polyA_signal
/gene="MEA6"
/note="first polyA signal"

```


2077	CTCTCTCTGACACTGGGTTTCTGTCACTCCATGGGACCGTAGGATGATGTTT	2136
2180	CTCCGCCAGGACAATCATATCTCTGATTCAGCCCTTCTCCACAAGGCCAAGACAGATTTT	2239
2137	CTCCGCCAGGACAATCATATCTCTGATTCAGCCCTTCTCCACAAGGCCAAGACAGATTTT	2196
2240	GTTCTAAATCTCGGTAGACTGTCCTGGACAGCAGAACTCAGAAGTTTTAAATGCTCTTT	2299
2197	GTTCTAAATCTCGGTAGACTGTCCTGGACAGCAGAACTCAGAAGTTTTAAATGCTCTTT	2256
2300	TGGATAAAATGGATGGGTCAATGCTTCAGAAATGGAAATCCAGTAGAAATGATACCAAG	2359
2257	TGGATAAAATGGATGGGTCAATGCTTCAGAAATGGAAATCCAGTAGAAATGATACCAAG	2316
2360	ATGATCTTTGGTAAATTTAAATGTGCTGATTCATCTCTCCCTGCTGAAATGAAGCCACTG	2419
2317	ATGATCTTTGGTAAATTTAAATGTGCTGATTCATCTCTCCCTGCTGAAATGAAGCCACTG	2376
2420	GCCCTGGCTTTTCTCTCCACTCTTGCTCCAAATCAGAGGTCCATGTTGTTCCAGTGGATG	2479
2377	GCCCTGGCTTTTCTCTCCACTCTTGCTCCAAATCAGAGGTCCATGTTGTTCCAGTGGATG	2436
2480	CAAGAGGCCCATTTCTTGAGAGAGACCTCTTTCCCCCCCACTCTCTCCAGGAGCAATG	2539
2437	CAAGAGGCCCATTTCTTGAGAGAGACCTCTTTCCCCCCCACTCTCTCCAGGAGCAATG	2496
2540	TTGGAGCTTCTCGAGATTAATTTCCACCAAGGGAATTCACAGGTCCACCACTGCTCCAT	2599
2497	TTGGAGCTTCTCGAGATTAATTTCCACCAAGGGAATTCACAGGTCCACCACTGCTCCAT	2556
2600	TTGCAATCAGAAATGCTATCCACGAGGGTTTTCTCTTTACCTTCCCCCAAGACCTG	2659
2557	TTGCAATCAGAAATGCTATCCACGAGGGTTTTCTCTTTACCTTCCCCCAAGACCTG	2616
2660	GATTTTTTCCCCACCCACCAATCTGAAGGTAGAAGTGAAGTTCCTCTCAGGTTTGATTC	2719
2617	GATTTTTTCCCCACCCACCAATCTGAAGGTAGAAGTGAAGTTCCTCTCAGGTTTGATTC	2676
2720	CACCTTCAAAATGAGCGTGTACTGAACATCCAGAACCCAGCAGAAAGAACTGACATATT	2779
2677	CACCTTCAAAATGAGCGTGTACTGAACATCCAGAACCCAGCAGAAAGAACTGACATATT	2736
2780	TTTGCTCTCTCCAAAGTAATTTTGACATGATCTCATTTTTCAGTTTAAAGTAACTGCTGTTA	2839
2737	TTTGCTCTCTCCAAAGTAATTTTGACATGATCTCATTTTTCAGTTTAAAGTAACTGCTGTTA	2796
2840	CTTAAGTGATTAACATTTTGCTCAAAATGCAAGCTTAATGGAATTAATTTCTCAGGATAG	2899
2797	CTTAAGTGATTAACATTTTGCTCAAAATGCAAGCTTAATGGAATTAATTTCTCAGGATAG	2856
2900	TATTTTGTAATTAAGATGATTTAAATATGAATCTTATCAGTAAATTTATTTCAATTTTAT	2959
2857	TATTTTGTAATTAAGATGATTTAAATATGAATCTTATCAGTAAATTTATTTCAATTTTAT	2916
2960	TTTAGACGGTATAACATATTTTCAATTTGATTAATCCACTAATTAATAAACAATAGTCGGAG	3019
2917	TTTAGACGGTATAACATATTTTCAATTTGATTAATCCACTAATTAATAAACAATAGTCGGAG	2976
3020	TTTTATATATGAATCTTTTCAGGTGGGGAGGCTTTAAATCTCGAAGTCTGTGCTTTATG	3079
2977	TTTTATATATGAATCTTTTCAGGTGGGGAGGCTTTAAATCTCGAAGTCTGTGCTTTATG	3036
3080	CCAGAACTGTATTTACTGTGTTGGACAATGTGGAAGTAACCTTATGCTTTAAATAA	3139
3037	CCAGAACTGTATTTACTGTGTTGGACAATGTGGAAGTAACCTTATGCTTTAAATAA	3096
3140	ATTATAGTTGATTTAAAAA	3158
3097	ATTATAGTTGATTTAAAGA	3115

RESULT 10

Db	18042	GTGGTATGTGCAGCTGTTATTTGGATTTTTTTGTGTCTCTCTTTTTTTGTGGAGAGTTTT	18101
Qy	541	AGATCGGTTAGGAGTCGGCTTTTACGTGGGAAGAGACAAAACCTTGTGTCAACGCTTTCT	600
Db	18102	AGATCGGTTAGGAGTCGGCTTTTACGTGGGAAGAGACAAAACCTTGTGTCAACGCTTTCT	18161
Qy	601	GGACTAATTGAAGAAAAATGTAACCTACTTTGAAAAATTTAGCCTTATTTCAAAAAGAGTAT	660
Db	18162	GGACTAATTGAAGAAAAATGTAACCTACTTTGAAAAATTTAGCCTTATTTCAAAAAGAGTAT	18221
Qy	661	GAAGGCTATGAAGTAGAGTCATCTTTAGAGTAGCCAGCTTTTGAGAGGAGCGCAGAGAA	720
Db	18222	GAAGGCTATGAAGTAGAGTCATCTTTAGAGTAGCCAGCTTTTGAGAGGAGCGCAGAGAA	18281
Qy	721	GAAGCAGCAAGTTTGGAGGCAACTCTGTGAAAAAGCTGAAACAGGTCCAAATCTTGAACCTTGA	780
Db	18282	GAAGCAGCAAGTTTGGAGGCAACTCTGTGAAAAAGCTGAAACAGGTCCAAATCTTGAACCTTGA	18341
Qy	781	GATGAAATCTCTGTCTTAGAAAAAGCTTTAAAAACAGAGAAATCTTAAACATTTCTCAACAA	840
Db	18342	GATGAAATCTCTGTCTTAGAAAAAGCTTTAAAAACAGAGAAATCTTAAACATTTCTCAACAA	18401
Qy	841	GATGAAATGATGCGCGATATTTCAAAAAGTATACAGTCTCTAGAAGATGAGTCAAAAATCC	900
Db	18402	GATGAAATGATGCGCGATATTTCAAAAAGTATACAGTCTCTAGAAGATGAGTCAAAAATCC	18461
Qy	901	CTCAAAATCAAAATAGCTGAGGCCAAAATCATCTCGACAGACATTTAAAAATGAGTGAAGAA	960
Db	18462	CTCAAAATCAAAATAGCTGAGGCCAAAATCATCTCGACAGACATTTAAAAATGAGTGAAGAA	18521
Qy	961	CGACGGGCTATAGCAATAAAAAGATGCTTTTGAATGAAAAATTTCTCAACTTTCAGACAAGCCAT	1020
Db	18522	CGACGGGCTATAGCAATAAAAAGATGCTTTTGAATGAAAAATTTCTCAACTTTCAGACAAGCCAT	18581
Qy	1021	AAACAGCTTTTTCAGCAAGAGCTGAACTATGGAAAGGAGAGTAGTGAGTGAACCTTAATAAA	1080
Db	18582	AAACAGCTTTTTCAGCAAGAGCTGAACTATGGAAAGGAGAGTAGTGAGTGAACCTTAATAAA	18641
Qy	1081	CAGAAAAATACATTTTGAAGACTCCTAAAGTACACGACAGAACAAAGTTCTGAAATGATAAGAA	1140
Db	18642	CAGAAAAATACATTTTGAAGACTCCTAAAGTACACGACAGAACAAAGTTCTGAAATGATAAGAA	18701
Qy	1141	AATCAGATCAAGACCTGACTGACACTTTGCCAATGATGAAGAGATCAGGCTGTGTGCTT	1200
Db	18702	AATCAGATCAAGACCTGACTGACACTTTGCCAATGATGAAGAGATCAGGCTGTGTGCTT	18761
Qy	1201	GAAGAACACACACGATGATATACCTGGAAATTAGAAGTGAACAGTGAATCGGAAAAAT	1260
Db	18762	GAAGAACACACATGGATGATGATTAACCTGGAAATTAGAAGTGAACAGTGAATCGGAAAAAT	18821
Qy	1261	GGTGCTTACTTAGATAATCCTCCAAAGGAGCTTTGAGAAACTGATTCATGCTGCTGAAG	1320
Db	18822	GGTGCTTAACTTAGATGATCCTCTGAAGGAGCTNTGAAGAACTGATTCATGCTGCTGAAG	18881
Qy	1321	TTAAATGCTTCTTTTAAAAACCTTTAGAAGAGAGAAAGAACCAAAATTTATATTCAGTTGTCT	1380
Db	18882	TTAAATGCTTCTTTTAAAAACCTTTAGAAGAGAGAAAGAACCAAAATTTATATTCAGTTGTCT	18941
Qy	1381	GAAGTTCATATAACAAAGGAGAGCTTACAGAGCATTTTAAAAATCTTCAGACTCAACAA	1440
Db	18942	GAAGTGCACAAAACAAAGGAAGAGCTTTACAGACATTTTAAAAATCTTCAGACTCAACAA	19001
Qy	1441	GCATCTTTTGAGTCAGAAAAACACATTTTGGAAAAATGAGAAATCAGAAGCTTCAACAGAAA	1500
Db	19002	GCATCTTTGCAATCAGAAAAACATATTTTGAAGGTGAGAAATCAGAAGCTTCAACAGAAA	19061
Qy	1501	CTTAAAGTAAATGACTGAATTTATATCAAGAAAAATGAATGAAATCCACAGGAAATTAACA	1560
Db	19062	CTTAAATAATGACTGAATTTCTATCAAGAAATGAATGAAATCTACAGGAAATTAACA	19121
Qy	1561	GTAGAGAAAAATTATCGGTTAGAGAAAGAGAAAACTTTCTTAAAGTAGATGAAAAAGATC	1620
Db	19122	GTGAGGAAAAATTACCGAATAGAGGAAGAGAGAGAGCTTTCTAGCTGGAAGGAAAGTCT	19181

Qy	1621	AGCCATGCCACTGAAGAGCTGGAGACCTATAGAAGCGAGCCAAAGATCTTTGAAGAGAA	1680
Db	19182	AGCCGTGCCACTGAACAGCTGGAGACCTATAGAAGCTAGCCAAAGATCTTTGAAGAGAA	19241
Qy	1681	TTGGAGAGAACTATTCAATCTTATCAAGGGCAGAGTATTTTCCCATGAGAAAAAGCACAT	1740
Db	19242	TTGGAGAGAACTGTTCAATTTTTATCAAAGCAGGTATTTCCTACGAGAAAAAGGAGCAT	19301
Qy	1741	GATAATTGTTGGCAGCTCGGAACTGCTGAAGAAACCTCAATGATTTTAAAGGAAGAAAT	1800
Db	19302	GATAATTGTTGGCAGCTCGGAACTGCTGAAGAAACCTCAATGATTTTAAAGGAAGAAAT	19361
Qy	1801	GCTCACAAACAGACAAAAATTAACTGAAACAGAGCTTAAATTTGAACTTTTAGAAAAAGAT	1860
Db	19362	GCTCACAAACAAACAAAAATTAACTGAGACAGAGTTGAAATTTGAACTTTTAGAAAAAGAT	19421
Qy	1861	CCTTATGCACTCGATGTTCCAAATACAGATTTGCGAGAGAGCATCCCCATATGCTCCC	1920
Db	19422	CCTTATGCACTCGATGTTTCAAAATACAGATTTGCGAGAGAGCATTTCCATGTAG--TCC	19479
Qy	1921	TCACCATTGGGTTGGCCCTTCATCTGAAAACAAGAGCTTTTCTCTCTCCCAACCTTTGTTG	1980
Db	19480	TCACCATTGGGTCGGCCCTTCATCTGAAAACGAGAGCTTTTCTCTCTCAAACTTTGTTG	19539
Qy	1981	GAGGTCCTACTCAGACTCTCACTTTGCTTCCAGGGGAGAGAGAGGCTCACGAGGC	2040
Db	19540	GAGGAT-CACTCAGACTCTCACTCTGTCTTCCAGGGGAGAGAGAGGCCCCAGCAGC	19598
Qy	2041	CCA-GGGAATCTCTGAGCCATCAGATTACCAATGAAAGAGAGGAATCAAGCTCTGATAG	2099
Db	19599	CCAGGGGATCCCTGACCATCAGATTACCAATGAAAGAGAGAGACCAGACTATGACAG	19658
Qy	2100	GTTAACCGATCTCATPAGAGCTCCCTCTGACACTGGGTCTGTGCACCTCCATGGGACCA	2159
Db	19659	GTTAATCGATCTCTCAGGGCTCTTCTGACACTGGGTC-CTGCTCATCTCCGGTGGAAACA	19717
Qy	2160	GGACCGTAGGATGATGTTTCTCCCGCCAGGACAAATCATCTGATTGAGCGCTTCCTCC	2219
Db	19718	GGACCGTAGGATGATGTTTCTCCACAGGGCAATCATATGCTGATTCAACTCTTCCTCC	19777
Qy	2220	ACAAAGCAGAGACAGATTTTGTCTAAATTCGTGTAGACTGTGTGACACAGAGAACTCAG	2279
Db	19778	ACAAAGGGAAGACAGATTTTATTTCTAAATCTGAAAGACTGTCTGGACAGCAGAAACCCAG	19837
Qy	2280	AAGTTTAAATATGCCCTCTTTGGATAAAATGGATGGGTCAATGCTTCAGAAATGGAAATC	2339
Db	19838	AAGTTTAAATAGCTTCTTTGGATAAAATGGATAGGTCAATGCTTCAGAAATGGAAATC	19897
Qy	2340	CAGTAGAAATGATACCAAGAGATGATCTTGGTAAATTTAAATGTGCTGATTCATCTCTCCC	2399
Db	19898	CAGTAGAAATGATGCCAAGAGATGATCTTGGTAAATTTAAATGTGCTGATTCATCTCTCCC	19957
Qy	2400	TGCTGAAAATGAAGCACTGGGCCCTGGCTTGTCTCCACTCTTGCTCCAAATCAGAGG	2459
Db	19958	TGCTGAAAATGAAGCAACTGGGCCCTGGCTTATTCTCTCCACTCTTGCTCCAGTACAGG	20017
Qy	2460	TCCATTGTTTCCAGTGGATGCAAGAGGCCCAATCTTTGAGAAGAGAGACCTCTTTCCGCC	2519
Db	20018	TCCATTGTTTCCAGTGGATACAG--GGGCGGTTCAAGAGAGAGAGACCTCTTTCCGCC	20076
Qy	2520	ACCTCTCCAGGAGCCATGTTTGGAGCTTCTCGAGATTATTTTCCACCAAGGATTTTCCC	2579
Db	20077	ACCTCTCTCCAGGAACCATGTTTGGAGCTTCTCGAGGTTATTTTCCACCAAGGATTTCCC	20136
Qy	2580	AGGTCCACCACTGCTCCATTGGCAATGAGAAATGCTATCCACCGAGGGGTTTTCCTCC	2639
Db	20137	AGGTCCACCACTGCTCCATTGGCAATGAGAAACATCTATCCACCGAGGGGTTTTCCTCC	20196
Qy	2640	TTACCTTTCCCCAAGACCTGGATTTTTTCCCCCAACCCCACTTCTGAAGGTAGAGTGA	2699
Db	20197	TTACCTTTCACTCCGAGACCTGGATTTT--ACCCTCAACCCCACTTCTGAAGGTAGAGCGA	20254

Db 87497 AGGTTAATTCATGAGAAAGACAGCAATTTTAAAAAGTGTATTTACCAAAAAATAACAATT 87438
QY 61 CAAATATCTGTATGTGAAACTGCCAGAACTTAAGCGCGCGGCTCAGACAGCGCTGC 120
Db 87437 CAAATATCTGTATGTGAAACTGCCAGAACTTAAGTGGTGGTTCGGACTCGCTGC 87378
QY 121 CTCAGGATGTAAGTGTAAACAAGAGGCGCAGGGGAGGTGGTGGGGAACAACATGGCGCTG 180
Db 87377 CCAGGATGTAAGTGTAAAGAAGGGGGCC- --GGGAGTGGTGGGGAACAACATGGCGCTG 87321
QY 181 TGAGGCGCTGTGGGTGCCCGCGTTCGCCAGCTCCCGCGAGCGCGCTCCACAGTGGTCCG 240
Db 87320 TGAGGCGCTGTGGGTGCCCGCGTTCGCCAGCTCCCGCGAGCGCGCTCCACAGCGTCCG 87261
QY 241 CTCGGTGTGGTGTCTACGTGCGCATTCGGGTCCAGACCCAAAGGTGGTGTCTCCACC 300
Db 87260 CTCGGTGTGGTGTCTACGTGCGCATTCGGGTCCAGACCCAAAGGTGGTGTCTCCACC 87201
QY 301 GCTGTGTGGCGAGTGTACTGCGGTGACCGCGCAGCAGCGCTCGAGCGTATGAGGAG 360
Db 87200 TTTTATGTGGCGAGTGTACTGCGGTGACCGCGCAGCAGCGCTCGGCGCTATGAGGAG 87141
QY 361 CTGCTGTGCTACCCCTCAGCGCTACCTGGGCGTGGTCTGGAGGAGCTACGACAGTGTG 420
Db 87140 CCGGTGTACCCCTCAGCGCTACCTGGGCGTGGTCTGGAGGAGCTACGACAGTGTG 87081
QY 421 GCAGCACTACCTCAGAGTATGAGACAGATGAGATCTTATGTTTCCATCGGAAGT 480
Db 87080 GCAGCACTACCTCAGAGTATGAGACAGATGAGATCTTATGTTTCCATCGGAAGT 87021
QY 481 GTGCTATGCGAGCTGTTATGGAATTTTGTGTTTCTCTTTTGTGGAGAGTGT 540
Db 87020 GTGCTATGCGAGCTGTTATGGAATTTTGTGTTTCTCTTTTGTGGAGAGTGT 86963
QY 541 AGATCGTTAGGAGTGGCTTTACGTGGGAAGAGAGCAAAACTTTGGTGCACGCTTCT 600
Db 86962 AGATCGTTAGGAGTGGCTTTACGTGGGAAGAGAGCAAAACTTTGGTGCACGCTTCT 86903
QY 601 GGAATAATTCAGAAAAAATGTAATCTACTTGAATAATTTAGCTTATTCAAAAAGAT 660
Db 86902 GGAATAATTCAGAAAAAATGTAATCTACTTGAATAATTTAGCTTATTCAAAAAGAT 86843
QY 661 GAAGCTATGAAGTAG- - - - -AGT 679
Db 86842 GAAGCTATGAAGTAGCTTTGAGAGTAATACTTTGAAAGTTTATGAGCTTATTCAAAA 86783
QY 680 CATCTTTAGAGGATGCCAGCTTTGAGAGGAGGAGCAGAGAGCAGAGCTTTGGAGG 739
Db 86782 GAGTATGAGGCGATGAAGTAGCTATGAGAGGCGCAGCAGAGCTTTGGAGG 86723
QY 740 CAACCTGTGAAGCTGAAACAGTCCATTTCTGAATTCGAGATGAATTCCTCTCTAG 799
Db 86722 CAACCTGTGAAGCTGAAACAGTCCATTTCTGAATTCGAGATGAATTCCTCTCTAG 86663
QY 800 AAAAGACTTTAAACAAGAGAAATCTAAACATTTCTCAACAAGATGAATTTGAGCGGATA 859
Db 86662 AAAAGACTTTAAACAAGAGAAATCTAAACATTTCTCAACAAGATGAATTTGAGCGGATA 86603
QY 860 TTTCAAAAAGTATACAGTCTCTAGAGATGAGTCAAAATCCCTCAAAATCAGAAATAGCTG 919
Db 86602 TTTCAAAAAGTATACAGTCTCTAGAGAGAGTCAAAATCCCTCAAAATCAGAAATAGCTG 86543
QY 920 AAGCCAAAATCATCTGCAAGACATTTAAATGAGTGAAGAGCGCGGCTATAGCAATAA 979
Db 86542 AAGCCAAAATCATCTGCAAGATTTTAAATGAGTGAAGAGCGCGGCTATAGCAATAA 86483
QY 980 AAGATCTTTGAATGAAATTTCTCAACTTCAGCAAGCCATAAACAGCTTTTTCAGCAAG 1039
Db 86482 AAGATCTTTGAATGAAATTTCTCAACTTCAGCAAGCCATAAACAGCTTTTTCAGCAAG 86426
QY 1040 AAGCTGAAGTGAAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1099
Db 86425 AAGCTGAAGTGAAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 86366

QY 1100 ACTCCAAAGTACACGAGCAACAAAGTTTCTGAATGATTAAGAAATACATCAAGACCTCA 1159
Db 86365 ACTCCAAAGTACACGAGCAACAAAGTTTCTGAATGATTAAGAAATACATCAAGACCTCA 86306
QY 1160 CTGGAACACTTTCGCAATGATGAAGATCAGCTGCTGTCTTGAAGAGACACAAACGATG 1219
Db 86305 CTGGAACACTTTCGCAATGATGAAGATCAGCTGCTGTCTTGAAGAGACACAAACGATG 86246
QY 1220 ATGATACCTGGAAATTTAGAGTGAACAGTGAATCGGAAATGCTGTCTTATAGATAATC 1279
Db 86245 ATGATACCTGGAAATTTAGAGTGAACAGTGAATCGG- - -AAATGGTGTAACTTAGATGATC 86187
QY 1280 CTCCAAAGAGGCTTTGAAAGAACTGATTCATGCTCTAAGTTTAAATGCTTTTAAAAA 1339
Db 86186 CTCCAAAGAGGCTTTGAAAGAACTGATTCATGCTCTAAGTTTAAATGCTTTTAAAAA 86127
QY 1340 CCTTTAGAGAGAGAAAGAAACCAAAATTTATATTCAGTTGCTGAAGTGTGATAAACAAGG 1399
Db 86126 GCTTTAGAGAGAGAAAGAAACCAATTTATTTAGTTTCTGAAATTTGATAAACAAGG 86067
QY 1400 AAGAGCTTACAGAGCATATTTAAAAATCTTCAGACTCAACAAGCATCTTTGCACTCAGAAA 1459
Db 86066 AAGAGCTTACAGAGCATATTTAAAAATCTTCAGACTCAACAAGCATCTTTGCACTCAGAAA 86007
QY 1460 ACACACATTTTGAAGATGAGAAATCAGAACTTCAACAGAACTTTAAAGTAATGACGTGAAT 1519
Db 86006 ACACACATTTTGAAGATGAGAAATCAGAACTTCAACAGAACTTTAAAGTAATGACGTGAAT 85947
QY 1520 TATATCAAGAAAAATGAAATGAAATCTCCAGGAAATTTAAACAGTAGAGGAAATTTATCGT 1579
Db 85946 TCTATCAAGAAAAATGAAATGAAATCTCCAGGAAATTTAAACAGTAGAGGAAATTTATCGG 85887
QY 1580 TAGAGAAAGAGAGAAACTTTCTAAAGTAGATGAAGATCAGCCATGCGCTGAGAGGC 1639
Db 85886 TAGAGAAAGAGAGAAACTTTCTAAAGTAGATGAAGATCAGCCATGCGCTGAGAGGC 85827
QY 1640 TGGAGAGCTTATGAAAGCGAGCCAAAGATCTTCAAGAGAAATTTGAGAGAACTTTCAIT 1699
Db 85826 TGGAGAGCTTATGAAAGCGAGCCAAAGATCTTCAAGAGAAATTTGAGAGAACTTTCAIT 85767
QY 1700 CTTATCAAGGCGAGATTTATTTCCATGAGAAAGACACATGATTAATTTGGTGGCAGCTC 1759
Db 85766 TTTATCAAAAGCAGGTTATTTCTTACGAGAAAGAGGACATGATTAATTTGGTGGCAGCTC 85707
QY 1760 GGAATGCTGAAAGAAACCTCAATGATTTAAGAAAGAAATGCTCAACAAGAGCAAAAAAT 1819
Db 85706 GGAATGCTGAAAGAAACCTCAATGATTTAAGAAAGAAATGCTCAACAAGAGCAAAAAAT 85647
QY 1820 TAACGAAAACAGAGCTTTAAATTTGAACTTTTGAAGAAAGATCCCTTATGCACTCGATGTT 1879
Db 85646 TAACGAAAACAGAGCTTTAAATTTGAACTTTTGAAGAAAGATCCCTTATGCACTCGATGTT 85587
QY 1880 CAAATACAGATTTGGCAGAGCATTCCCATATGCTCCCTCACCATTGGGTTGGCCTT 1939
Db 85586 CAAATACAGATTTGGCAGAGCATTCCCATATGCTCCCTCACCATTGGGTTGGCCTT 85527
QY 1940 CATCTGAAAACAGAGCTTTTCTCTCTCCCTCAACTTTTGTGGAGGTTCCACTCAGACTCT 1999
Db 85526 CATCTGAAAACAGAGCTTTTCT 85467
QY 2000 CACCTTTGCTTCCAGGGGAGGAGGAGGCTCAGAGGCCAGGGAATCTCTCTGAGCC 2059
Db 85466 GACCTGTCTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 85407
QY 2060 ATCAGATTAACAATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2119
Db 85406 ATCAGATTAACAATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 85347
QY 2120 CTCCTCTGACACTGGGCT 2179
Db 85346 CTCCTCTGACACTGGGCT 85287

QY	2180	CTCCGCCAGGACAAATCATATCTGATTCAGCCCTTCTCTCCACAAAGGCAAGACAGATTTT	2239
Db	85286	CTCCACACGAGCAAAATCATATCTGATTCAGCTCTTCTCTCTCAAAGGAAAGACAGATTTT	85227
QY	2240	GTTCTAAATCTGGTAGACTGTCTGGACAGCAGAACTCAGAAGTTTTAAATATGCCTTCTT	2299
Db	85226	ATTCTAAATCTGAAAGACATGTCTGGACAGCAGAAACCCAGAAGTTTTAAATGACTTCTT	85167
QY	2300	TGGATAAAATGGATGGGTCAATGCCCTTCAGAAATCGAATCCAGTAGAATAATGATACCAAAG	2359
Db	85166	TGGATAAAATGGATGGGTCAATGCCCTTCAGAAATCGAATCCAGTAGAATAATGATCCCAAAG	85107
QY	2360	ATGATCTTGGTAAATTTAAATATGGCTGATTTCATCTCTCCCTGCTGAAATGAAGCCACTG	2419
Db	85106	ATGATCTTGGTAAATTTAAATATGGCTGATTTCATCTCTCCCTGCTGAAATGAAGCAACTG	85047
QY	2420	GCCCTGGCTTTGTTCTCTCAACCTCTTCTCTCAATCAGAGGTCCATTTGTTCCAGTGGATG	2479
Db	85046	GCCCTGGCTTTATTCCTCTCACCTCTTGCTCCAGTCAGAGGCCAATTTGTTTCCAGTGGATA	84987
QY	2480	CAAGAGGCCCATTTCTTGAAGAAGAGCACTTCCTTTCCGCCCACTCTCTCCAGGAGCCATGT	2539
Db	84986	CAAGGGGCCGTTTCATGAGAAGAGGATCTCTTTCCGCCCACTCTCTTCAGGAACCATGT	84927
QY	2540	TTGGAGCTTCTCGAGATTATTTTCCACCAAGGGATTTCCAGGTCCACACCTGCTCCTCAT	2599
Db	84926	TTGGAGCTTCTCGAGGTATTTTCCACCAAGGGATTTCCAGGTCCACACCATGCTCCTCAT	84867
QY	2600	TTGCAATGAGAAATGCTATCCACGAGGGGTTTTCCCTCTTACCTTCCCCCAAGACACTG	2659
Db	84866	TTGCAATGAGAACAATCATACCCACCAAGGGTTTACCTCTTACCTTCATCCGAGACTG	84807
QY	2660	GAATTTTCCCCCACCACCAATCTGAAGGTAGAAGTAGTTCCTCCCTCAGGTTTGATTC	2719
Db	84806	GAATTTT - ACCCCCAACCCACACATCTGAAGGTAGAAGCGAGTTCCTCTCAGGATTGATTC	84749
QY	2720	CACCTTCAAATGAGCTGCTACTGAACATCCAGAACCAAGCAAGAAACCTTGACAAATAT	2779
Db	84748	CGCTTTCAA - - - - -TGCTACTGAATCCAGGCCACAGCAAGAAACCTTGACAAATAT	84696
QY	2780	TTTGCTCTCTTCAAAAGTAATTTTGACTGATCTCATTTTTCAGTTTAAGTAACCTGCTGTTA	2839
Db	84695	GTGCTCTCTCTCAAAAGTAATTTTGACGGATCTCATTTTCAGTTTAAGTAACCTGCTATTA	84636
QY	2840	CT - - - - -TAAGTGATTACACTTTTGTCTCAAATTTGAAGCTTAATGGAATTTAATTTCTCAGG	2895
Db	84635	CTTAAGTAAAGTGAATTCACATTTTGTCTCAAATTTGAAGTTTAATGGAATTTAATTTCTCAGG	84576
QY	2896	ATAGTATTTTGTAAATAAGATGATTTAAATATGAATCTTATGAGTAATAATTTATTTCAATT	2955
Db	84575	ATAGTATTTTGTAAATAAGATGCTTTAAATATGAATCTTATGAGTAATAATTTATTTCAATT	84516
QY	2956	- - -TTATTTTAGCGGTATAACTATTTCATTTTG - - -ATTAACTCACTATTATTATAAACA	3009
Db	84515	TATTAATCTAGATCGTATAACTATTATTAAATTTTGGTGAATTAATCCACTGTTATAGAAACA	84456
QY	3010	ATAGTCGGAGTTTTTATATATGTAATCTTTTCAGGTGGGAGGCTTTAAATTTCTGAAGTCT -	3068
Db	84455	ATATTTGGGAGTTTTTATATATGTAATCTTTCAGGTGGGAGGCTTTAAATTTCTGAAGGTTG	84396
QY	3069	- -GTGTCTTTATGCCAAGAACTGATTTTACTGTGGTTGTGGACAAATGTGAAAGTAACCT	3126
Db	84395	TGCTGTCTTCATGCCAAGAACTGTATTCACTGTGGTTGTAGATAAATGTGAAAGTAACCT	84336
QY	3127	TATGCTTTAAATAAATTTATAGTTGATTTA 3154	
Db	84335	TATGCTTTAAATTTAATAAACTTTTAGTTGA 84308	

RESULT 12
HSM807156
LOCUS HSM807156 2931 bp mRNA linear PRI 28-AUG-2003
DEFINITION Homo sapiens mRNA; cDNA DKFZP686F08252 (from clone DKFZP686F08252);

complete cds.
BX640994 BX640994.1 GI:34365350
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2931)
Lauber, J., Bahr, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.
The German Human cDNA Consortium
Direct Submission
Submitted (26-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email: s.wiemann@kfz-heidelberg.de; sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFPz686F08252) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://mips.gsf.de/proj/cDNA/
Location/Qualifiers
1..2931
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="14q21.1"
/clone="DKFPz686F08252"
/tissue_type="human salivary gland"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host DH10B_sites SfIIA + SfiIB"
/dev_stage="adult"
1..2931
/gene="DKFPz686F08252"
310..2484
/gene="DKFPz686F08252"
/notes="meningiona-expressed antigen 6 (MEAG), differentially spliced"
/codon_start=1
/product="hypothetical protein"
/protein_id="CA545997.1"
/db_xref="GI:34365351"
/tranlation="MLSGLIEEKQLKFLPSLVQKEGYEVSESSLKDASFEKEATEA QSLAEATCKLNSRSLBDEMLCLEKLEKSKSHQEDELMDADISKIQSLDESKS LKSOVAEKWTFKIFOMNERKIKATKDALNENSOIQESOKQLQAEVYKEQVSELN XQVTFEDSKVHAQVLYNDEKSHIKTLTERLLMKMDAAMLCAGEDITDDONLELVNSE SENGATIDPPKAGVLYNDEKSHIKTLTERLLMKMDAAMLCAGEDITDDONLELVNSE LQTKQASLSENTHFNENKQKQKQKVTLEYQENMKHLKTLVTEENRLEKBEKRL SKYDEKISHATELETRYRAKOLEBELRTHSYQGIISHEKHADNWLAAARNAER NINDLAKENAHNRKQITELPEKEDPYALDVDPNTAFGRHSGPYGSPGLWGNSE TRAFLEPPTLLLEQPLRLSPLLPGGGRSGRGNPLDHOITNBERGSSCDRLTDPHRA PDSPTGSLSPWDDRRMPPPGQSYSDSALPORODRSCNSGRSLSGPAELFSFNP SLDMKGSFSPSENESSRNDTKDLGNLVPDSLPAENATGCGFVPPPLAPRGDPLF VDMARGFLRRGFPFPFPFGAGRSRDYFPFRDPFGFPFPAPFAMRVYPPRGFPFPY LPFRPGFPFPFPFSEGRSEFPFSGSLIPSPNEPATHEHPPOQET"
2890..2895
polyA_signal
/gene="DKFPz686F08252"
2910
polyA_site
/gene="DKFPz686F08252"
ORIGIN
Query Match 82.0%; Score 2593.8; DB 9; Length 2931;
Best local Similarity 97.3%; Pred. No. 0;
Matches 2671; Conservative 0; Mismatches 67; Indels 7; Gaps 3;
Qy 415 GTTGTGCACACTACTGAGATATGAGACCAGATGAGATCCTATGCTTTCATCG 474
Db 132 GTTGTGCACACTGCCTGAAGATATGAGACCAGATCTTAATTTTATGCTTTCATCG 191
Qy 475 GAACCTGGTATATGTCAGCTGTTATTTGGATTTTTTGTCTTCTC-CTTTTTTGTGGAG 533

Db 192 GAAATTGGTGATATGTGACACTGTTGTGGATTTTGTGCTCTTTTGTGGAG 251
QY 534 AAGTTTATAGTCGGTTAGAGTCGGCTTTACGTGGGAAGAGACAAAACCTTGTGCAAC 593
Db 252 AAGTTTATAGTCGGTTAGAGTCGGCTTTATGTGGGACGAGAAAAAGCTTGTCTAAT 311
QY 594 GCTTTCTGACATAATTGAAGAAAAATGTAAACTACTTGTAAAAATTTAGCCTTATTCAAA 653
Db 312 GCTTCTGACATAATTGAAGAAAAATTAATACTACTTGAANAATTTAGCCTTGTTCAAA 371
QY 654 AGAGTATGAAGCTATGAAGTATGATCATCTTTAGAGATGCCAGCTTTGAGAGGAGGC 713
Db 372 AGAGTATGAAGCTATGAAGTATGATCATCTTTAAAGATGCCAGCTTTGAGAGGAGGC 431
QY 714 ASCAGAGAGACGACGAAGTTTGGAGCACTGTGAAGAGCTGACAGCTCCATTCGA 773
Db 432 AAC - - - AGAAGCACAAAGTTTGGAGCAACCTTGTGAAGAGCTGAACAGCTCCATTCGA 488
QY 774 ACTTGAGGATGAATCCTCTGTCTAGAAAAAGACTTTAAACAGAGAAATCTTAAACATTC 833
Db 489 ACTTGAGGATGAATCCTCTGTCTAGAAAAAGAGTTTAAAGAGAGAAATCCAAACATTC 548
QY 834 TCACAGAGATGAATGTATGGCGGATATTCAAAAGTATACAGTCTCTAGAGATGAGTC 893
Db 549 TGAACAGATGAATTTGATGGCGGATATTCAAAAGGATACAGTCTCTAGAGATGAGTC 608
QY 894 AAAATCCCTCAATCACAAATAGCTGAAGCCAAAATCATCTGCAAGACATTTTAAATGAG 953
Db 609 AAAATCCCTCAATCACAAATAGCTGAAGCCAAAATGACCTTCAAGATATTTCAATGAA 668
QY 954 TGAAGAACAGCGGCTATAGCAATAAAGATGCTTTGAATGAANAATCTCAACITTCAGAC 1013
Db 669 TGAAGAACAGCTGAAGATAGCAATAAAGATGCTTTGAATGAANAATCTCAACITTCAGGA 728
QY 1014 AAGCCATAAAGACTTTTTCAGCAAGAGCTGAAAGTATGAAAGAGAGAGTGAAGTGAAC 1073
Db 729 AAGCCAGAACAGCTTTT - - - GCAAGAGCTGAGTATGAAAGAACAGTGAAGTGAAC 785
QY 1074 TAATAACAGAAAAATTAACATTTGAAGACTCCAAAGTACAGCAGAACAAAGTCTGAATGA 1133
Db 786 TAATAACAGAAAAATTAACATTTGAAGACTCCAAAGTACAGCAGAACAAAGTCTGAATGA 845
QY 1134 TAAAGAAATACATCAAGACCTGACTGGAACACTTGCCAAATGATGAAGAACAGGCTGC 1193
Db 846 TAAAGAAATGATCAATCAAGACTGACTGAAAGCTTTTAAAGTGAAGATTTGGCTGC 905
QY 1194 TGTGCTTGAAGAGACACAAACGATGATGAATCAACCTGGAATTTAGAGTGAACAGTGAATC 1253
Db 906 TATGCTTGAAGAGACATAACCGATGATGAATCACTTGAATTTAGAAATGAACAGTGAATC 965
QY 1254 GGAATAATGCTTACTTAGATAATCTCCAAAGAGGCTTTTGAAGAACTGATTCATGC 1313
Db 966 GGAATAATGCTTACTTAGATAATCTCCAAAGAGGCTTTTGAAGAACTGATTCATGC 1025
QY 1314 TGCTAAGTAAATGCTTCTTTTAAACCTTTAGAGGAGAGAAACCAAAATTTATTTCA 1373
Db 1026 TGCTAAGTAAATGCTTCTTTTAAACCTTTAGAGGAGAGAAACCAAAATTTATTTCA 1085
QY 1374 GTTGTCTGAAGTTGATAAACAAGAGAGCTTACAGAGCATATTTAAATAATCTTCAGAC 1433
Db 1086 GTTGTCTGAAGTTGATAAACAAGAGAGCTTACAGAGCATATTTAAATAATCTTCAGAC 1145
QY 1434 TCACAGAGCTTTTGCAGTCAGAAACACACATTTTGAATAATGAATCAGAACTTCA 1493
Db 1146 TCACAGAGCTTTTGCAGTCAGAAACACACATTTTGAATAATGAATCAGAACTTCA 1205
QY 1494 ACAGAACTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1553
Db 1206 ACAGAACTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1265
QY 1554 ATTAACAGTAAAGAAAAATATCGGTTAGAGAAAGAGAAACCTTTTAAAGTATGATGA 1613

Db 1266 ATTAACAGTAAAGAAAAATATCGGTTAGAGAAAGAGAGAACCTTTCTTAAAGTATGA 1325
QY 1614 AAGATCAGCATGCGCACTGAAGAGCTGGAGACCTTATAGAAAGCGAGCCAAAGATCTTGA 1673
Db 1326 AAGATCAGCATGCGCACTGAAGAGCTGGAGACCTTATAGAAAGCGAGCCAAAGATCTTGA 1385
QY 1674 AGAAGAAATGGAGAGAACTATTCAATCTTATCAAGGGCAGATTTATTTCCCATGAGAAAA 1733
Db 1386 AGAAGAAATGGAGAGAACTATTCAATCTTATCAAGGGCAGATTTATTTCCCATGAGAAAA 1445
QY 1734 AGCAGATGATTAATTTGGTTGGCAGCTCGGAATGCTGAAAGAAACCTCAATGATTTAAGGA 1793
Db 1446 AGCAGATGATTAATTTGGTTGGCAGCTCGGAATGCTGAAAGAAACCTCAATGATTTAAGGA 1505
QY 1794 AGAAATGCTTCACACACAGACAAAAATAAAGTGAACAGAGCTTAAATTTGAACCTTTTGA 1853
Db 1506 AGAAATGCTTCACACACAGACAAAAATAAAGTGAACAGAGCTTAAATTTGAACCTTTTGA 1565
QY 1854 AAAAGATCTTATGCACTGATGTTCCAAATACAGATTTGGCAGAGAGATTTCCCATTA 1913
Db 1566 AAAAGATCTTATGCACTGATGTTCCAAATACAGATTTGGCAGAGAGATTTCCCATTA 1625
QY 1914 TGGTCCCTCACCATTGGGTTGGCTTTCATCTGAAACAAGAGCTTTTCTCTCTCTCCAAC 1973
Db 1626 TGGTCCCTCACCATTGGGTTGGCTTTCATCTGAAACAAGAGCTTTTCTCTCTCTCCAAC 1685
QY 1974 TTTGTTGGAGGTCCTCAGACTCTCACTTTGCTTCCAGGGGAGAGAGAGAGGCTC 2033
Db 1686 TTTGTTGGAGGTCCTCAGACTCTCACTTTGCTTCCAGGGGAGAGAGAGAGGCTC 1745
QY 2034 ACAGAGGCCAGGGAATCCTCTGAGACCATCAGATTTACCAATGAAAGAGAGAGATCAAGCTG 2093
Db 1746 ACAGAGGCCAGGGAATCCTCTGAGACCATCAGATTTACCAATGAAAGAGAGAGATCAAGCTG 1805
QY 2094 TGATAGTTTAAACCGATCTCTCAGAGCTCCCTCTGACACTGGGTCTCTGTCACCTCCATG 2153
Db 1806 TGATAGTTTAAACCGATCTCTCAGAGCTCCCTCTGACACTGGGTCTCTGTCACCTCCATG 1865
QY 2154 GGACAGGACCGTAGAGATGATTTCTCCGCGCAGAGCAATCATATCTGATTCAGCCCT 2213
Db 1866 GGACAGGACCGTAGAGATGATTTCTCCGCGCAGAGCAATCATATCTGATTCAGCCCT 1925
QY 2214 TCCTCCACAAAGGCAAGACAGATTTTGTCTAATTTCTGGTAGACTGTCTGGACAGAGAGA 2273
Db 1926 TCCTCCACAAAGGCAAGACAGATTTTGTCTAATTTCTGGTAGACTGTCTGGACAGAGAGA 1985
QY 2274 ACTCAGAGATTTAATATGCTCTTTTGGATAAATGGATGGGTCAATGCTCTTCAGAAAT 2333
Db 1986 ACTCAGAGATTTAATATGCTCTTTTGGATAAATGGATGGGTCAATGCTCTTCAGAAAT 2045
QY 2334 GGAATCCAGTAGAAATGATACCAAGATGATCTTGGTAATTTTAAATGTCCTGATTCATC 2393
Db 2046 GGAATCCAGTAGAAATGATACCAAGATGATCTTGGTAATTTTAAATGTCCTGATTCATC 2105
QY 2394 TCCTCCCTGCTGAAATGAAGCACTGGCCTGGCTTTGTTCTCCACCTCTTCTGCTCCAAT 2453
Db 2106 TCCTCCCTGCTGAAATGAAGCACTGGCCTGGCTTTGTTCTCCACCTCTTCTGCTCCAAT 2165
QY 2454 CAGAGGTCCATTTGTTTCCAGTGGATGCAAGAGGCCCATTTCTTGAGAGAGAGACCTCCTTT 2513
Db 2166 CAGAGGTCCATTTGTTTCCAGTGGATGCAAGAGGCCCATTTCTTGAGAGAGAGACCTCCTTT 2225
QY 2514 CCCCCACCTCTCCAGAGCCATGTTTGGAGCTTCTCGAGATTTATTTCCACAGAGGA 2573
Db 2226 CCCCCACCTCTCCAGAGCCATGTTTGGAGCTTCTCGAGATTTATTTCCACAGAGGA 2285
QY 2574 TTTCCAGGTCCACCACTCTCATTTGCAATGAGAAATGTCTATCCACCGAGGGGTTT 2633
Db 2286 TTTCCAGGTCCACCACTCTCATTTGCAATGAGAAATGTCTATCCACCGAGGGGTTT 2345
QY 2634 TCCTCTTTA CTTTCCCCCAAGACCTGGAATTTTCCCCCAGCCCCCAGATTTCTGAAGTAG 2693
Db 2346 TCCTCTTTA CTTTCCCCCAAGACCTGGAATTTTCCCCCAGCCCCCAGATTTCTGAAGTAG 2405


```

QY 2694 AAGTGAGTCCCTCAGGTTTGATTCACCTTCAATGAGCCTGCTACTGACATCCAGA 2753
Db 2406 AAGTGAGTCCCTCAGGTTTGATTCACCTTCAATGAGCCTGCTACTGACATCCAGA 2465
QY 2754 ACCACGCAAGAACTGACATATTTTGGCTCTTCAAAAGTAAATTTGACTGATCTC 2813
Db 2466 ACCACGCAAGAACTGACATATTTTGGCTCTTCAAAAGTAAATTTGACTGATCTC 2525
QY 2814 ATTTTCAGTTTAAAGTAACTGCTGTTACTTAAAGTGATTAACATTTTGCTCAAAATGAAGCT 2873
Db 2526 ATTTTCAGTTTAAAGTAACTGCTGTTACTTAAAGTGATTAACATTTTGCTCAAAATGAAGCT 2585
QY 2874 TAATGAATATTAATTTCTCAGGATAGTATTTTGTAAATAAGATGATTTAAATATGAATC 2933
Db 2586 TAATGAATATTAATTTCTCAGGATAGTATTTTGTAAATAAGATGATTTAAATATGAATC 2645
QY 2934 TTATGAGTAAATTAATTTCAATTTTATTTAGACGGTATAACATTTTCAATTTGATTAATC 2993
Db 2646 TTATGAGTAAATTAATTTCAATTTTATTTAGACGGTATAACATTTTCAATTTGATTAATC 2705
QY 2994 CACTATTATTAACATAGTGGAGTTTATATATATGTAATCTTTCAGTGGGAGGCTT 3053
Db 2706 CACTATTATTAACATAGTGGAGTTTATATATGTAATCTTTCAGTGGGAGGCTT 2765
QY 3054 TAAATCTGAGTCTGTCTTTATGCGCAAGAACTGTATTTACTGTGGTTGTGGACAAAT 3113
Db 2766 TAAATCTGAGTCTGTCTTTATGCGCAAGAACTGTATTTACTGTGGTTGTGGACAAAT 2825
QY 3114 GTGAAAGTAACTTTATGCTTAAATAATAATATAGTTGATTTAAAAA 3158
Db 2826 GTGAAAGTAACTTTATGCTTAAATAATAATATAGTTGATTTAAAAA 2870

RESULT 13
HSM0808160 2957 bp mRNA linear PRI 30-AUG-2003
LOCUS Homo sapiens mRNA; cDNA DKF2p686M2159 (from clone DKF2p686M2159).
DEFINITION BX648014
ACCESSION BX648014
VERSION BX648014.1 GI:34367173
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2957)
Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
Wewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
The German Human cDNA Consortium
Direct Submission
Submitted (27-AUG-2003) MIPs, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKF2p686M2159) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.ssf.de/proj/cDNA/.
Location/Qualifiers
1..2957
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKF2p686M2159"
/tissue_type="human endometrium"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host
DH10B; sites SfiI + SfiIB"
/dev_stage="adult"
FEATURES
source

```

```

polyA_signal 2846..2851
polyA_site 2864
ORIGIN
Query Match 81.6%; Score 2583; DB 9; Length 2957;
Best Local Similarity 97.2%; Pred. No. 0;
Matches 2673; Conservative 0; Mismatches 70; Indels 8; Gaps 4;
QY 415 GTTGTGCGAGCACTACTGAGAGTATGAGACGAGATGAGAAATCCTTATGCTTTCCATCG 474
Db 132 GTTGTGCGAGCACTGCGCTGAAGGTATGAGACGAGATTTCTAATCTTTATGCTTTCCATCG 191
QY 475 GAACTGTGGTATCTGAGCTGTTTATGAGATTTTGTGCTCTC-CTTTTGTGCGAG 533
Db 192 GAATTTGTGATGTTGAGCTGTTTGTGATTTTGTGCTCTCTTTTGTGCGAG 251
QY 534 AAGTTTATAGTCGTTTAGGAGTCGGCTTTTACGTGGGAAGAGACCAAAATCTTGCTGCAAC 593
Db 252 AAGTTTATAGTCGTTTAGGAGTCGGCTTTTATGTTGGACGAGAGAAAAGCTTGCTCTAAT 311
QY 594 GCTTTCGGACTAATTCGAAGAAAATGTAACACTACTTTGAAAATTTTACGCTTATTCAAA 653
Db 312 GCTTTCGGACTAATTCGAAGAAAATGTAACACTACTTTGAAAATTTTACGCTTATTCAAA 371
QY 654 AGAGTATGAAGGCTATGAAGTAGAGTCATCTTTAGAGATGCCAGCTTTGAGAGAGGCG 713
Db 372 AGAGTATGAAGGCTATGAAGTAGAGTCATCTTTAAGGATGCCAGCTTTGAGAGAGGCG 431
QY 714 AGCAGAGACGACGAGTTTGGAGGCAACCTGTGAAAAGCTGAACAGGTCCTCAATTCGA 773
Db 432 AAC---AGAGACACAAGTTTGGAGGCAACCTGTGAAAAGCTGAACAGGTCCTCAATTCGA 488
QY 774 ACTTGAGGATGAAATCCTCTCTCTAGAAAAGAGCTTTAAAACAGAGAAATCTAAACATTC 833
Db 489 ACTTGAGGATGAAATCCTCTCTCTAGAAAAGAGCTTTAAAACAGAGAAATCTAAACATTC 548
QY 834 TCACAGAGCAATGATGCGGGATATTTCAAAAGTATACAGTCTCTAGAGATGAGTC 893
Db 549 TGAACAAGCACTGAACATAGCAATAAAAGATGCTTTGAAATGAAATTTCTCAACTTCAGA 728
QY 1014 AAGCCATAAAACAGCTTTTTCAGCAAGAGCTGAAAGTATGGAAGGAGAGAGTGAAGTGA 1073
Db 729 AAGCCAGAAACAGCTTTT---GCAAGAGCTGAAAGTATGGAAGGAGAGCAAGTGAAGTGA 785
QY 1074 TAATAAAGCAAGAAATAACATTTTGAAGATCTCAAAAGTACACCCAGAACAGTCTGAATGA 1133
Db 786 TAATAAAGCAAGAAATAACATTTTGAAGATCTCAAAAGTACATCAGAACAGTCTTAATGA 845
QY 1134 TAAGAAATACATCAAGACCTTCTGACCTGGACACTTGCCTCAATGATGAAGATCAGGCTGC 1193
Db 846 TAAGAAATACATCAAGACCTTCTGACCTGGACCTTGTAAAGATGAAGATTCGGCTGC 905
QY 1194 TGTGCTTGAAGAGACACCAACGATGATGATACTCGAATAGAGTGAAGTGAACAGTGAATC 1253
Db 906 TATGCTTGGAGAAGACATAACCGATGATGATACTTGAATTTAGAAATGAACAGTGAATC 965
QY 1254 GGAAATGCTGCTTACTTAGATAATCTCCAAAAGGAGCTTTGAAAGAACTGATTCATGC 1313
Db 966 GGAAATGCTGCTTACTTAGATAATCTCCAAAAGGAGCTTTGAAAGAACTGATTCATGC 1025
QY 1314 TGCTAAGTTAAATGCTTCTTTAAAAACCTTAGAAGAGAGAAACCAAAATTTATATTC 1373
Db 1026 TGCTAAGTTAAATGCTTCTTTAAAAACCTTAGAAGAGAGAAACCAAAATTTATATTC 1085
QY 1374 GTTGCTGAGTGTGATAAACAAGAGAGGCTTACAGAGCATATTTAAAAATCTTCAGAC 1433

```

1086 GTTGTCTGAAGTTGATATAAAACAAAGAGAGCTTACAGAGCATATTAATAATCTTCAGAC 1145
 1434 TCAACAAGCATCTTTGCGAGTCAGAAAACACACATTTTGAAAATCAGAAATCAGAAAGCTTCA 1493
 1146 TGAACAGACATCTTTGCGAGTCAGAAAACACACATTTTGAAAATCAGAAATCAGAAAGCTTCA 1205
 1494 ACAGAAAATTAAAGTAATGACTGAATATATCAAGAAAATGAATGAATGAAATCTCCACAGAA 1553
 1206 ACAGAACTTTAAAGTAATGACTGAATATATCAAGAAAATGAATGAATGAAATCTCCACAGAA 1265
 1554 ATTAAACAGTAGAGGAAAATTTATCGGTTAGAGAAAAGAGAAAATCTTTAAAGTAGATGA 1613
 1266 ATTAAACAGTAGAGGAAAATTTATCGGTTAGAGAAAAGAGAAAATCTTTAAAGTAGATGA 1325
 1614 AAAGATCAGCCATGCCACTGAAGAGCTGAGACCTATAGAAGCGAGGCAAGATCTTTGA 1673
 1326 AAAGATCAGCCATGCCACTGAAGAGCTGAGACCTATAGAAGCGAGGCAAGATCTTTGA 1385
 1674 AGAAGAAATGGAGAGAACTATTCTTCTTATCAAGGGCAGAGTTATTTCCCATGAGAAAAA 1733
 1386 AGAAGAAATGGAGAGAACTATTCTTCTTATCAAGGGCAGAGTTATTTCCCATGAGAAAAA 1445
 1734 AGCACAATGATAATTTGGTGGCAGCTCGGAATGCTGAAAGAAACCTCAATGATTTAAGGAA 1793
 1446 AGCACAATGATAATTTGGTGGCAGCTCGGAATGCTGAAAGAAACCTCAATGATTTAAGGAA 1505
 1794 AGAAGATGCTCAACAGACAAAAATTAATCTGAAACAGAGCTTAAATTTGAATTTTGA 1853
 1506 AGAAGATGCTCAACAGACAAAAATTAATCTGAAACAGAGCTTAAATTTGAATTTTGA 1565
 1854 AAAAGATCTTTATGCACTCGATGTTCCAAAATACAGATTTGGCAGAGAGCAATCCCCATA 1913
 1566 AAAAGATCTTTATGCACTCGATGTTCCAAAATACAGATTTGGCAGAGAGCAATCCCCATA 1625
 1914 TGGTCCCTCAGCANTTTGGTGGCTTCATCTGAAACAGAGCTTTTCTCTCTCTCCCAAC 1973
 1626 TGGTCCCTCAGCANTTTGGTGGCTTCATCTGAAACAGAGCTTTTCTCTCTCTCCCAAC 1685
 1974 TTTGTTGGAGGGTCCACTCAGACTCTCAGCTTTGCTTCCAGGGGAGGAGGAGGCTC 2033
 1686 TTTGTTGGAGGGTCCACTCAGACTCTCAGCTTTGCTTCCAGGGGAGGAGGAGGCTC 1745
 2034 ACAGAGCCAGGAAATCTCTGACCATCAGATTAACATGAAGAGGAGATCAGCTG 2093
 1746 ACAGAGCCAGGAAATCTCTGACCATCAGATTAACATGAAGAGGAGATCAGCTG 1805
 2094 TGATAGTTAAACCGATCCTCATAGGCTCCCTCTGACACTGGGCTCTCTGCTCACCTCCATG 2153
 1806 TGATAGTTAAACCGATCCTCATAGGCTCCCTCTGACACTGGGCTCTCTGCTCACCTCCATG 1865
 2154 GGACCCAGGACCGTAGGATGATTTGCTCGCCAGGACATCATATCTCTGATTCAGCCCT 2213
 1866 GGACCCAGGACCGTAGGATGATTTGCTCGCCAGGACATCATATCTCTGATTCAGCCCT 1925
 2214 TCCTCCCAAGGCAAGACAGATTTGTTCTTAATCTGAGACTGTCTGGACAGCAGAGA 2273
 1926 TCCTCCCAAGGCAAGACAGATTTGTTCTTAATCTGAGACTGTCTGGACAGCAGAGA 1985
 2274 ACTCAGAAAGTTTAAATGCTCTCTTTGGATTAAGATGAGTGGGTCAATGCTTCAGAAAT 2333
 1986 ACTCAGAAAGTTTAAATGCTCTCTTTGGATTAAGATGAGTGGGTCAATGCTTCAGAAAT 2045
 2334 GGAATCAGTAGAATATGATACCAAGATGATCTTTGTAATTTAAATGCTGCTGATTCATC 2393
 2046 GGAATCAGTAGAATATGATACCAAGATGATCTTTGTAATTTAAATGCTGCTGATTCATC 2105
 2394 TCTCCCTGCTGAAAATGAAGCACTGCGCCCTGCTTTGTTTCTCCCACTCTTCTGCTCAAT 2453
 2106 TCTCCCTGCTGAAAATGAAGCACTGCGCCCTGCTTTGTTTCTCCCACTCTTCTGCTCAAT 2165
 2454 CAGAGGTCCTATGTTTCCAGTGGATGCAAGAGGCCCATCTCTGAGAGAGAGCTCTCTTT 2513

2166 CAGAGGTCAATGTTTCCAGTGGATGCAAGAGGCCCATCTCTTGAGAGAGGAGCCTCTTT 2225
 2514 CCCCCACCTCTCCAGAGGCCATGTTTGGAGCTTCTCGAGATTATTTTCCACCAAGGA 2573
 2226 CCCCCACCTCTCCAGAGGCCATGTTTGGAGCTTCTCGAGATTATTTTCCACCAAGGA 2285
 2574 TTTCCAGAGTCCACCACTGCTCCATTTGCAATGAGAAATGCTATCCACCCAGGAGGTTT 2633
 2286 TTTCCAGAGTCCACCACTGCTCCATTTGCAATGAGAAATGCTATCCACCCAGGAGGTTT 2345
 2634 TCTCTCTTACCTTCCCCCAAGACCTGGATTTTCCCCCAAGACCTGCTGCTGCTGCTGCT 2693
 2346 TCTCTCTTACCTTCCCCCAAGACCTGGATTTTCCCCCAAGACCTGCTGCTGCTGCTGCT 2405
 2694 AGGTGAGTTCCCTCAGGTTTGGATTCACCTTCAATGAGCCTGCTGCTGCTGCTGCTGCT 2753
 2406 AGGTGAGTTCCCTCAGGTTTGGATTCACCTTCAATGAGCCTGCTGCTGCTGCTGCTGCT 2465
 2754 ACCACAGAGAAACCTGACATATTTTGGTCTCTCTTCAAAAGTAAATTTGACTGATCTC 2813
 2466 ACCACAGAGAAACCTGACATATTTTGGTCTCTCTTCAAAAGTAAATTTGACTGATCTC 2524
 2814 ATTTTCAGTTTAAAGTAACTGCTGCTTAAAGTAACTGCTGCTGCTGCTGCTGCTGCTGCT 2873
 2525 ATTTTCAGTTTAAAGTAACTGCTGCTTAAAGTAACTGCTGCTGCTGCTGCTGCTGCTGCT 2584
 2874 TAATGGAATTTAATTTCTCAGGATAGTATTTTGAATAAAGATGATTTAAATATGAAATC 2933
 2585 TAATGGAATTTAATTTCTCAGGATAGTATTTTGAATAAAGATGATTTAAATATGAAATC 2644
 2934 TTATGAGTAAATTTTCAATTTTATTTAGACGGTATTAATTTCAATTTTCAATTTGATTAATC 2993
 2645 TTATGAGTAAATTTTCAATTTTATTTAGACGGTATTAATTTCAATTTTCAATTTGATTAATC 2704
 2994 CACTATTATATAAACAATAGTGGAGTTTATATATGTAATCTTTCAAGTGGGAGGCTT 3053
 2705 CACTATTATATAAACAATAGTGGAGTTTATATATGTAATCTTTCAAGTGGGAGGCTT 2764
 3054 TAAATTTGAGTCTGCTGCTTTATGCAAGAGTCTGTTTACTGCTGCTGCTGCTGCTGCTGCT 3113
 2765 TAAATTTGAGTCTGCTGCTTTATGCAAGAGTCTGTTTACTGCTGCTGCTGCTGCTGCTGCT 2824
 3114 GTGAAAGTAACTTTATGCTTTAAATAATATAGTATTTTAAATAAAAAA 3164
 2825 GTGAAAGTAACTTTATGCTTTAAATAATATAGTATTTTAAATAAAAAA 2875

RESULT 14
 HSM803540
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 HSM803540
 Homo sapiens mRNA; cDNA DKFP686F2453 (from clone DKFP686F2453).
 AL832233
 AL832233.1 GI:21732779
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 2890)
 Wambutt, R., Heubner, D., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
 Pobo, G., Han, M. and Wiemann, S.
 Direct Submission
 Submitted (13-MAY-2003) MIPS, Ingolstaedter Landstr.1, D-85764
 Neuherberg, GERMANY
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
 sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
 consortium of the German Genome Project.
 This clone (DKFP686F2453) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further
 information about the clone and the sequencing project is available
 at http://mips.gsf.de/proj/cDNA/.

FEATURES

Location/Qualifiers

1..2890
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp686f2453"
 /clone_type="human endometrium"
 /clone_lib="686 (synonym: hicc3). Vector pSport1_sfi; host DH10B; sites SfiIA + SfiIB"
 /dev_stage="adult"
 2850..2855
 polyA_signal
 2870
 polyA_site
 ORIGIN

Query Match 80.0%; Score 2531.8; DB 9; Length 2890;

Best Local Similarity 96.3%; Pred. No. 0;

Matches 2644; Conservative 0; Mismatches 72; Indels 29; Gaps 4;

QY	415	GTGTGGCGACACTACCTGAGAGTATGAGACAGATGAGAACTTATGTGTTCCATCG	474
DB	114	GTGTGGCGACACTGCTGAAGTATGAGACAGATCTTAATCTTTATGGTTTCATGG	173
QY	475	GAATCGTGTGTGTGAGCTGTTATTTGGATTTTTTGTCTCTCTTTTTTGTGGAGA	534
DB	174	GAATCGTGTGTGTGAGCTGTTATTTGGATTTTTTGTCTCTCTTTTTTGTGGAGA	233
QY	535	AGTTTGTAGTCGTTAGGAGTCGGCTTACGTGGGAGAGAGCAAAACTTGGTCCACG	594
DB	234	AGTTTGTAGTCGTTAGGAGTCGGCTTACGTGGGAGAGAGCAAAACTTGGTCCATG	271
QY	595	CTTTCTGGACTAATTGAAGAAAAATTAACCTACTTTGAAAAATTTAGCCCTTATCAAAA	654
DB	272	CTTTCTGGACTAATTGAAGAAAAATTAACCTACTTTGAAAAATTTAGCCCTTATCAAAA	331
QY	655	GAGTATGAAGGCTATGAGTATGAGTATCTTTAGAGATGCCAGCTTTGAGAGAGGCA	714
DB	332	GAGTATGAAGGCTATGAGTATGAGTATCTTTAGAGATGCCAGCTTTGAGAGAGGCA	391
QY	715	GCAGAGAGACACGAGTTTGGAGGCAACCTGTGAAAAAGCTGAACAGGTCCTAATCTGAA	774
DB	392	AC--AGAGACACAAAGTTTGGAGGCAACCTGTGAAAAAGCTGAACAGGTCCTAATCTGAA	448
QY	775	CTTGAGATGAATCTCTCTCTAGAAAAAGACTTAAACAGAGAAATCAAACTTCT	834
DB	449	CTTGAGATGAATCTCTCTCTAGAAAAAGACTTAAACAGAGAAATCAAACTTCT	508
QY	835	CAACAGATGAATGATGGCGGATTTTCAAAAGATATACAGTCTCTAGAGATGAGTCA	894
DB	509	GAACAGATGAATGATGGCGGATTTTCAAAAGATATACAGTCTCTAGAGATGAGTCA	568
QY	895	AAATCCCTCAATCAAAATAGCTGAAGCAAAATCATCTGCAAGACATTTAAATGAGT	954
DB	569	AAATCCCTCAATCAAAATAGCTGAAGCAAAATCATCTGCAAGACATTTAAATGAGT	628
QY	955	GAAGACGCGGCTATAGCAATAAAGATGCTTTGAATGAAATTTCACTTCTAGACA	1014
DB	629	GAAGACGCTGAACATAGCAATAAAGATGCTTTGAATGAAATTTCACTTCTAGACA	888
QY	1015	AGCCATTAACAGCTTTTTCAGCAAGAGAGCTGAAGTATGGAAGAGAGAGTGAAGCTT	1074
DB	689	AGCCATTAACAGCTTTTTCAGCAAGAGAGAGCTGAAGTATGGAAGAGAGAGTGAAGCTT	745
QY	1075	AATTAACAGAAATTAACATTTGAAGCTCCAAAGTACACGAGCAAGTCTCAATGAT	1134
DB	746	AATTAACAGAAATTAACATTTGAAGCTCCAAAGTACATGAGCAAGTCTCAATGAT	805
QY	1135	AAAGAAATCAATCAAGACCTGACTGGACACTTGGCAATGATGAAGATCAGGCTGCT	1194
DB	806	AAAGAAATCAATCAAGACCTGACTGAGCAAGCTTTTAAAGATGAAGATTTGGCTGCT	865
QY	1195	GTGCTTGAAGAGACACAAACGATGATGATCACTGGAATTAGAAGTGAACAGTGAATCG	1254
DB	866	ATGCTTGAAGAGACATTAACGATGATGATCACTTGAATTTAGAAATGAACAGTGAATCG	925

QY	1255	GAATAATGGTCTTACTTAGATAATCTCCAAAAAGAGCTTTGAAGAACTGATTCTATGCT	1314
DB	926	GAATAATGGTCTTACTTAGATAATCTCCAAAAAGAGCTTTGAAGAACTGATTCTATGCT	985
QY	1315	GCTAAGTTAAATGCTCTTTTAAAAAACCCTTAGAGAGGAGAAAGAACCAAAATTTATATT	1374
DB	986	GCTAAGTTAAATGCTCTTTTAAAAAACCCTTAGAGAGGAGAAAGAACCAAAATTTATATT	1045
QY	1375	TTGTCTGAAGTTGATAAAAAAGAGAGAGCTTACAGAGCATATTTAAAAATCTTCAGACT	1434
DB	1046	TTGTCTGAAGTTGATAAAAAAGAGAGAGCTTACAGAGCATATTTAAAAATCTTCAGACT	1105
QY	1435	CAACAAGCATTTTGCAGTCAGAAAAACACACATTTTGAATAAGAGATCAGAAGCTTCAA	1494
DB	1106	GAACAAGCATTTTGCAGTCAGAAAAACACACATTTTGAATAAGAGATCAGAAGCTTCAA	1165
QY	1495	CAGAAACTTAAAGTAATGACTGAAATATATCAAGAAAAATGAAATCTCCACAGAAA	1554
DB	1166	CAGAAACTTAAAGTAATGACTGAAATATATCAAGAAAAATGAAATCTCCACAGAAA	1225
QY	1555	TTAACAGTAGAGGAAAAATTTATCGGTTAGAGAAAGAGAGAACTTTCTAAAGTAGATGAA	1614
DB	1226	TTAACAGTAGAGGAAAAATTTATCGGTTAGAGAAAGAGAGAACTTTCTAAAGTAGATGAA	1285
QY	1615	AAGATCAGCCATGCCACTGAAGAGCTGAGACCTATAGAAAGCGAGCCAAAGATCTTGA	1674
DB	1286	AAGATCAGCCATGCCACTGAAGAGCTGAGACCTATAGAAAGCGAGCCAAAGATCTTGA	1345
QY	1675	GAAAGTTGGAGAGAACTATTCATTCTTATCAAGGCGAGATTTTCCCTGAGAAAAA	1734
DB	1346	GAAAGTTGGAGAGAACTATTCATTCTTATCAAGGCGAGATTTTCCCTGAGAAAAA	1405
QY	1735	GCATCATATAATTTGGTGGCAGCTCGGAATGCTGAAAGAACTCAATGATTTAAGAAA	1794
DB	1406	GCATCATATAATTTGGTGGCAGCTCGGAATGCTGAAAGAACTCAATGATTTAAGAAA	1465
QY	1795	GAAAGTTGCTCAACAGAGAAATTAATCTGAACAGAGCTTAAATTTGAATTTAGAA	1854
DB	1466	GAAAGTTGCTCAACAGAGAAATTAATCTGAACAGAGCTTAAATTTGAATTTAGAA	1525
QY	1855	AAAGATCCTTATGACCTCGATGTTCCAAATACAGATTTGGCAGAGAGCAATCCCAATAT	1914
DB	1526	AAAGATCCTTATGACCTCGATGTTCCAAATACAGATTTGGCAGAGAGCAATCCCAATAT	1585
QY	1915	GGTCCCTCACATTTGGTGGCCTTCATCTGAACAGAGCTTTTCTCTCTCTCCCACT	1974
DB	1586	GGTCCCTCACATTTGGTGGCCTTCATCTGAACAGAGCTTTTCTCTCTCTCCCACT	1645
QY	1975	TTGTTGGAGGTTCCACTCAGACTCTCACTTTGCTTCCAGGGGAGAGAGAGGCTCA	2034
DB	1646	TTGTTGGAGGTTCCACTCAGACTCTCACTTTGCTTCCAGGGGAGAGAGAGGCTCA	1705
QY	2035	CGAGCCCGAGGAAATCTCTCGACCATCAGATTTACCAATGAAGAGAGAAATCAAGCTGT	2094
DB	1706	CGAGCCCGAGGAAATCTCTCGACCATCAGATTTACCAATGAAGAGAGAAATCAAGCTGT	1765
QY	2095	GATAGTTAAACGATCTCTAGAGCTCCCTCTGACACTGGGTCTCTGTCACTCCATGG	2154
DB	1766	GATAGTTAAACGATCTCTAGAGCTCCCTCTGACACTGGGTCTCTGTCACTCCATGG	1825
QY	2155	GACCAAGACCTGAGATGATGTTTCTCCGCGAGAGCAATCATATCTGATTCAGCCCTT	2214
DB	1826	GACCAAGACCTGAGATGATGTTTCTCCGCGAGAGCAATCATATCTGATTCAGCCCTT	1885
QY	2215	CTCCACAAAGCAGACAGATTTTGTCTTAATCTGCTAGACTGCTTGGAGCAGAGAA	2274
DB	1886	CTCCACAAAGCAGACAGATTTTGTCTTAATCTGCTAGACTGCTTGGAGCAGAGAA	1945
QY	2275	CTCAGAGTTTAAATATGCTTCTTTTGGATAAAAATGGATGGGTCAATGCTTCAAGATG	2334
DB	1946	CTCAGAGTTTAAATATGCTTCTTTTGGATAAAAATGGATGGGTCAATGCTTCAAGATG	2005

2335 GATCCAGTAAATGATACCAAGATGATCTTGGTAAATTTAAATGCTGCTGATTCATCT 2394
2395 GATCCAGTAAATGATACCAAGATGATCTTGGTAAATTTAAATGCTGCTGATTCATCT 2065
2395 CTCCTGCTGAAATGAAGCCACTGCGCTGCTTGTCTCCACCTCTTCTCCCAATC 2454
2066 CTCCTGCTGAAATGAAGCCACTGCGCTGCTTGTCTCCACCTCTTCTCCCAATC 2125
2455 AGAGTCCATGCTTCCAGTGAATGAGAGCCCATCTTCTGAGAGGAGGACCTCTTTC 2514
2126 AGAGTCCATGCTTCCAGTGAATGAGAGCCCATCTTCTGAGAGGAGGACCTCTTTC 2185
2515 CCCCCACCTCTCCAGGAGCCATGTTGAGAGCTTCTCGAGATTTATTTTCCACCAAGGAT 2574
2186 CCCCCACCTCTCCAGGAGCCATGTTGAGAGCTTCTCGAGATTTATTTTCCACCAAGGAT 2245
2575 TTCCAGGTCACCACTCTGCTCAATTTGCAATGAGAAATGCTATCTCCACGAGGGTTT 2634
2246 TTCCAGGTCACCACTCTGCTCAATTTGCAATGAGAAATGCTATCTCCACGAGGGTTT 2305
2635 CTTCTTACCTTCCCCCAAGACTGATTTTCCCCCACCACCCACATTTCTGAAGGTAGA 2694
2306 CTTCTTACCTTCCCCCAAGACTGATTTTCCCCCACCACCCACATTTCTGAAGGTAGA 2365
2695 AGTGAG-TTCCCTCAGGTTGATTTCCACCTTCAATGAGCTGCTACTGACATCCAGA 2753
2366 AGTGAGTTTCCCTCAGGTTGATTTCCACCTTCAATGAGCTGCTACTGACATCCAGA 2425
2754 ACCACAGCAAGAACTGCAATATTTTGGCTCTCTCAAAAGTAAATTTGACTGATCTC 2813
2426 ACCACAGCAAGAACTGCAATATTTTGGCTCTCTCAAAAGTAAATTTGACTGATCTC 2485
2814 ATTTTCAGTTTAAAGTAACTGCTTAAAGTAACTGCTTAAAGTAACTGCTTAAAGTAACT 2873
2485 ATTTTCAGTTTAAAGTAACTGCTTAAAGTAACTGCTTAAAGTAACTGCTTAAAGTAACT 2545
2874 TAATGAATTAATTTCTCAGGATAGTATTTTGTAAATTAAGATGATTTTAAATATGAATC 2933
2546 TAATGAATTAATTTCTCAGGATAGTATTTTGTAAATTAAGATGATTTTAAATATGAATC 2605
2934 TTATGAGTAAATTAATTTTAAATTTTAAAGGATTAATTAATTTTAAATTTTAAATTTTAAAT 2993
2606 TTATGAGTAAATTAATTTTAAATTTTAAAGGATTAATTAATTTTAAATTTTAAATTTTAAAT 2665
2994 CACTATTAATAACATAGTGGAGTTTATATATGTAATCTTTTCAAGTGGGAGGCTT 3053
2666 CACTATTAATAACATAGTGGAGTTTATATATGTAATCTTTTCAAGTGGGAGGCTT 2725
3054 TAAATTTGAGGCTGCTGCTTTTATGCCAAGACTGATTTTAAATTTTAAATTTTAAATTTTAAAT 3113
2726 TAAATTTGAGGCTGCTGCTTTTATGCCAAGACTGATTTTAAATTTTAAATTTTAAATTTTAAAT 2785
3114 GTCAAGTAACTTTATGCTTAAATTAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 3158
2786 GTCAAGTAACTTTATGCTTAAATTAATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 2830

RESULT 15
BC064355
LOCUS BC064355
DEFINITION Homo sapiens cDNA clone MGC:75562 IMAGE:4803802, complete cds.
ACCESSION BC064355
VERSION BC064355.1 GI:39963693
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2895)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schmitt, T.E., Brownstein, M.J., Udén, T.B., Teshi, Yuki, S.,
Carninci, P., Frange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Lu, X., Gibbs, R.A.,
Worley, K.K., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Pahey, J., Helton, E., Kettman, M., Madan, A., Young, A.C., Rodriguez, S.,
Sanchez, A., Whiting, M., Touchman, J.W., Green, E.D.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E.,
Schnerk, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2895)
Strausberg, R.
Direct Submission
Submitted (15-DEC-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov
Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Grante, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masiello, C., Maekri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripoop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, O.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 134 Row: n Column: 8
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.

FEATURES
source
1. 2895
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:75562 IMAGE:4803802"
/tissue_type="Skin, normal"
/clone_lib="NCI CGAP_Skn3"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
283..2712
/codon_start=1
/product="Unknown (protein for MGC:75562)"
/protein_id="AAH64355.1"
/db_xref="GI:39963693"

CDS

LYQENMKLRKLTVEENVYRLEKEKLSKVDEKISHATELETVKRAKDLEELERT
IH5YQOIIISHEKSHKADHNWLAARNAERLNDLRKENAHNRQKLTETELKFLLEKDPY
ALDVENTAETREHSPYGLWPSSETAFPLSPLEGLPURLSPLLPFGGGRSGRG
PQNLHDQITNERSGSSCDRLDHPHSDTGLSPDMDQDRMMFPPPGSYDPSAL
PQRODRFCNSGRSLGPAELRSFNMPLSLDKMDGMPSEMESRMDTDXDLGNLVPD
SSLPABENATGFGFVPPPLAPVRGFLFVVDARGPFLRRGPPPPPPGAMFGASRDYF
PGDFPFPAPAPAMRNVPVPPRGFPFLYPPPPRGFPFPFPHPHSGRSEFFPSGLIPSPNEP
ATEHPPEQOET"
misc_feature
1309..1749
/notes="APG6; Region: Autophagy protein Apg6. In yeast, 15
Apg proteins coordinate the formation of autophagosomes.
Autophagy is a bulk degradation process induced by
starvation in eukaryotic cells. Apg6/Vps30 has two
distinct functions in the autophagic process, either
associated with the membrane or in a retrieval step of the
carboxypeptidase Y sorting pathway"
/db_xref="CDD:pfam04111"

ORIGIN

Query Match 80.0%; Score 2531.4; DB 9; Length 2895;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 2702; Conservative 0; Mismatches 136; Indels 39; Gaps 5;

QY 90 CTAAGGGGGCCGGCTCAGACAGCGCTGCTCAGAGTGTAAAGTGAACAGAGGGCC 149
DB 25 CTTGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 84
QY 150 AGGGAGGTGGTGGGGGCAACATGGGCTGTGAGGCTGTGGGTGGGTGGGTGGGTGG 209
DB 85 GGGATGG-----CAGGTAGGCTGTGAGGCTGTGGGTGGGTGGGTGGGTGGGT 134
QY 210 CTCGCCCGGAGCCCGCTCAGAGTGTGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 269
DB 135 CTCGCCCGGAGCCCGCTCAGAGTGTGCTCCGCTCCGCTCCGCTCCGCTCCGCTCC 190
QY 270 GTTCCAGACCAAGGCTGGTGTCTCCAGCGTGTGTGT-----GGCGAGTGT 319
DB 191 GTTCCAGACCAAGGCTGGTGTCTCCAGCGTGTGTGT-----GGCGAGTGT 250
QY 320 ACTCGGTGACCGGACAGAGCGCTCGAGTGTGAGAGCGCTGTGTACCCCTCAGC 379
DB 251 ACTGTGCGACCAAGAGAGCGCTTTGGCGCTATGAGAGCGCGCGGCTTACCCCTCAAC 310
QY 380 CCTACCTGGGGTGGTCTCGAGAGCTAGCAGAGTGTGGCAGCAGTACCTGAGAGTA 439
DB 311 CGGATTTGGGGTGGTCTCGAGAGCTAGCAGAGTGTGGCAGCAGTACCTGAGAGTA 370
QY 440 TGAGACAGATGAGATCCTTATGGTTTCCATCGGAACCTGGTGTATGTCAGCTGTTA 499
DB 371 TGAGACAGATCCTTATGGTTTCCATCGGAATTTGGTGTATGTCAGCTGTG 430
QY 500 TTGATTTTGTGTTCTCTCTTTTGTGGAGAGTTTGTAGTCGGTTAGGATCGGC 559
DB 431 TTGATTTTGTGTTCTCTCTTTTGTGGAGAGTTTGTAGTCGGTTAGGATCGGC 490
QY 560 TTTACGTGGGAGAGAGCAAAACCTTGGTGCACGCTTCTGGACTAATTTGAAGAAAAAT 619
DB 491 TTTATGTGGGACGAGAGAAAAAGCTTCTCTATGCTTTCTGGACTAATTTGAAGAAAAA 550
QY 620 GTAACCTACTTGAAAAATTAGCTTATCMAAAGAGTATGAGGCTATGAGTAGAGT 679
DB 551 GTAACCTACTTGAAAAATTAGCTTATCMAAAGAGTATGAGGCTATGAGTAGAGT 610
QY 680 CATCTTTAGAGGATGCCAGCTTTGAGAGGAGGAGGAGCAGAGAA-----GAAGCAC 727
DB 611 CATCTTTAAGGATGCCAGCTTTGAGAGGAGGAGCAGAGAACAAAGTTTGGAGGTAG 670
QY 728 GAAGTTTGGAGGCACTGTGAAAGCTGAACAGGTCCAATCTGAACTTGAGATGAAA 787
DB 671 AAAATCAAATGGCACTGTGAAAGCTGAACAGGTCCAATCTGAACTTGAGATGAAA 730
QY 788 TCCTCTGTCTAGAAAAAGACTTAAAAAAGAGAGTCTAAACATTTCTCAACAGATGAAT 847

DB 731 TACTCTCTAGAAAAAGAGCTTAAAGAGAGAGAAATCCAAACATTCTTGAACAAGATGAAT 790
QY 848 TGATGGCGGATATTTCAAAGATATACAGTCTCTAGAGATGAGTCAAAATCCCTCAAT 907
DB 791 TGATGGCGGATATTTCAAAGAGATACAGTCTCTTAGAAGATGAGTCAAAATCCCTCAAT 850
QY 908 CACAAATAGCTGAAGCCAAATCATCTGCAAGACATTTAAATAGAGTGAAGAGCAGCGG 967
DB 851 CACAGTAGCTGAAGCCAAATGACCTTCAAGATATTTCAATGAATGAAGAGCAGCTGA 910
QY 968 CTATAGCAATAAAGATGCTTTGAATGAAATTTCTCACTTACAGCAAGCATTAACAGC 1027
DB 911 ACATAGCAATAAAGATGCTTTGAATGAAATTTCTCACTTACAGCAAGCATTAACAGC 970
QY 1028 TTTTTCAGCAAGAGCTGAAAGTATGGAAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGA 1087
DB 971 TTTT---GCAAGAGCTGAGATGGAAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 1027
QY 1088 TAACTTTGAAGACTCCAAAGTACACCGCAGAAACAGTCTGAATGATGAAGAAATCACA 1147
DB 1028 TAACATTTGAAGACTCCAAAGTACATGCAACAAAGTCTTAAATGATGAAGAAATCACA 1087
QY 1148 TCAAGACCTGACCTGGACACTTGCCTAATGATGAAGATCAGGCTGCTGCTTGAAGAAG 1207
DB 1088 TCAAGCTCTGACCTGAACGCTTGTAAAGATGAAGATTTGGGCTGCTATGCTTGGAGAG 1147
QY 1208 ACACAACCGATGATGAACCTGGAATTAGAAGTGAACAGTGAAGTGAAGTGAAGTGAAG 1267
DB 1148 ACATAACCGATGATGAACCTTGAATTAGAATGAACAGTGAAGTGAAGTGAAGTGAAG 1207
QY 1268 ACTTAGATAATCTTCCAAAGAGCTTTGAGAGAACTGATTCATGCTGCTTAAGTAAATG 1327
DB 1208 ACTTAGATAATCTTCCAAAGAGCTTTGAGAGAACTGATTCATGCTGCTTAAGTAAATG 1267
QY 1328 CTTCTTTAAAAACCTTAGAAGGAGAGAAACCAAAATTTATATTCAGTGTGCTGAAAGTTG 1387
DB 1268 CTTCTTTAAAAACCTTAGAAGGAGAGAAACCAAAATTTATATTCAGTGTGCTGAAAGTTG 1327
QY 1388 ATAAACAGAGAGAGAGCTTACAGAGCATATTAAGAACTTTCAGACTCAACAGCATCTT 1447
DB 1328 ATAAACAGAGAGAGAGCTTACAGAGCATATTAAGAACTTTCAGACTCAACAGCATCTT 1387
QY 1448 TGCAGTCAGAAAAACACACATTTTGAAGATGAGAAATCAGAACTTCAACAGAAATCTTAAAG 1507
DB 1388 TGCAGTCAGAAAAACACACATTTTGAAGATGAGAAATCAGAACTTCAACAGAAATCTTAAAG 1447
QY 1508 TAATGACTGAATTTATCAAGAAATGAATGAACCTCCAGAGAAATTAACAGTAGAGG 1567
DB 1448 TAATGACTGAATTTATCAAGAAATGAATGAACCTCCAGAGAAATTAACAGTAGAGG 1507
QY 1568 AAAATTATCGGTAGAGAGAGAGAACTTTCTAAAGTAGATGAAAGATCAGCCATG 1627
DB 1508 AAAATTATCGGTAGAGAGAGAGAGAACTTTCTAAAGTAGATGAAAGATCAGCCATG 1567
QY 1628 CCACTGAAGAGCTGAGAGCTTATGAAGAGCGAGCCAAAGATCTTGAAGAAATTTGGAGA 1687
DB 1568 CCACTGAAGAGCTGAGAGCTTATGAAGAGCGAGCCAAAGATCTTGAAGAAATTTGGAGA 1627
QY 1688 GAACATTTCTTCTTATCAAGGGCAGATTTATTTCCATGAGAAAAAGACATGATAAT 1747
DB 1628 GAACATTTCTTCTTATCAAGGGCAGATTTATTTCCATGAGAAAAAGACATGATAAT 1687
QY 1748 GGTTCGAGCTCGGAATGCTGAAAGAAACCTCAATGATTTAAGGAAAGAAATGCTCA 1807
DB 1688 GGTTCGAGCTCGGAATGCTGAAAGAAACCTCAATGATTTAAGGAAAGAAATGCTCA 1747
QY 1808 ACAGACAAAAATTAACCTGAAACAGAGCTTAAATTTGAACCTTTTGAAGAAAGATCCTTATG 1867
DB 1748 ACAGACAAAAATTAACCTGAAACAGAGCTTAAATTTGAACCTTTTGAAGAAAGATCCTTATG 1807
QY 1868 CACTCGATGTTTCCAAATACAGCATTTGGCAGAGCATTTCCCATATGCTGCTCCCAT 1927
DB 1808 CACTCGATGTTTCCAAATACAGCATTTGGCAGAGCATTTCCCATATGCTGCTCCCAT 1867

QY 1928 TGGGTTGGCCCTCATCTGAAACAGAGAGCTTTCTCTCTCCCTCCAACTTTGTTGGAGGGTC 1987
Db |||||
QY 1868 TGGGTTGGCCCTCATCTGAAACAGAGAGCTTTCTCTCTCCCTCCAACTTTGTTGGAGGGTC 1927
Db |||||
QY 1988 CACTCAGACTCTCAGCTTTGCTTCCAGGGGAGAGAGGCTCAGAGGCCACAGGA 2047
Db |||||
QY 1928 CACTCAGACTCTCAGCTTTGCTTCCAGGGGAGAGAGGCTCAGAGGCCACAGGA 1987
Db |||||
QY 2048 ATCCTCTGGACATCAGATTAACCAATGAAGAGGAGATCAAGCTGTGATAGGTTAACCG 2107
Db |||||
QY 1988 ATCCTCTGGACATCAGATTAACCAATGAAGAGGAGATCAAGCTGTGATAGGTTAACCG 2047
Db |||||
QY 2108 ATCCTCATAGGGCTCCCTCTGACACTGGGTCTCTGTACCTCCATGGGACCCAGACCGTA 2167
Db |||||
QY 2048 ATCCTCATAGGGCTCCCTCTGACACTGGGTCTCTGTACCTCCATGGGACCCAGACCGTA 2107
Db |||||
QY 2168 GGATGATGTTTCCCTCCGACGAGACATCATATCTGTATTCAGCCCTTCTCCCAAAAGGC 2227
Db |||||
QY 2108 GGATGATGTTTCCCTCCGACGAGACATCATATCTGTATTCAGCCCTTCTCCCAAAAGGC 2167
Db |||||
QY 2228 AAGACAGATTTTGTCTAATTCGTAGACTGTCTGGACCCAGCAGAACTCAGAAAGTTTAA 2287
Db |||||
QY 2168 AAGACAGATTTTGTCTAATTCGTAGACTGTCTGGACCCAGCAGAACTCAGAAAGTTTAA 2227
Db |||||
QY 2288 ATATCCCTTCTTTGGATTAATGAGATGGGTCTATGCCCTTCAGAAATGGAATCCAGTAGAA 2347
Db |||||
QY 2228 ATATCCCTTCTTTGGATTAATGAGATGGGTCTATGCCCTTCAGAAATGGAATCCAGTAGAA 2287
Db |||||
QY 2348 ATGATACCAAGATGATCTTGGTAAATTTAAATGTGCTGATTCATCTCTCCCTGCTGAAA 2407
Db |||||
QY 2288 ATGATACCAAGATGATCTTGGTAAATTTAAATGTGCTGATTCATCTCTCCCTGCTGAAA 2347
Db |||||
QY 2408 ATGAAGCCACTGGCCCTGGCTTTGCTTCTCAGCTCTTGTCTCCAAATCAGAGGTCCATTGT 2467
Db |||||
QY 2348 ATGAAGCCACTGGCCCTGGCTTTGCTTCTCAGCTCTTGTCTCCAAATCAGAGGTCCATTGT 2407
Db |||||
QY 2468 TTCCAGTGATGCAAGAGGCCCATTTCTGAGAGAGGACCTCTTCCCTCCACCTCCTC 2527
Db |||||
QY 2408 TTCCAGTGATGCAAGAGGCCCATTTCTTGGAGAGAGGACCTCTTCCCTCCACCTCCTC 2467
Db |||||
QY 2528 CAGGAGCCATGTTGGAGCTTCTGAGATTAATTTCCACAGGGGATTTCCAGGTCCAC 2587
Db |||||
QY 2468 CAGGAGCCATGTTGGAGCTTCTGAGATTAATTTCCACAGGGGATTTCCAGGTCCAC 2527
Db |||||
QY 2588 CACCTGCTCCATTTGCAATGAGAAATGCTATCCACCGAGGGGTTTCTCTTACCTTC 2647
Db |||||
QY 2528 CACCTGCTCCATTTGCAATGAGAAATGCTATCCACCGAGGGGTTTCTCTTACCTTC 2587
Db |||||
QY 2648 CCCCAGACCTGGATTTTCCCCCACCACCTCTCTGAAGGTAGAAGTTCCTCCCT 2707
Db |||||
QY 2588 CCCCAGACCTGGATTTTCCCCCACCACCTCTCTGAAGGTAGAAGTTCCTCCCT 2647
Db |||||
QY 2708 CAGGTTGATTCACCTTCAATGAGCTGCTACTGACATCCAGACCCAGACCCAGAGAA 2767
Db |||||
QY 2648 CAGGTTGATTCACCTTCAATGAGCTGCTACTGACATCCAGACCCAGACCCAGAGAA 2707
Db |||||
QY 2768 CTTGACATATTTTCTCTCTTCAAAAGTAAATTTGACTGATCTCATTTTCAGTTTAAAG 2827
Db |||||
QY 2708 CTTGACATATTTTCTCTCTTCAAAAGTAAATTTGACTGATCTCATTTTCAGTTTAAAG 2767
Db |||||
QY 2828 TAACTGCTGTTACTTAAGTATTAAGTATTTCTCTCAATTTGAAGCTTAATGGAATTATAA 2887
Db |||||
QY 2768 TAACTGCTGTTACTTAAGTATTAAGTATTTCTCTCAATTTGAAGCTTAATGGAATTATAA 2827
Db |||||
QY 2888 TTCTCAGGATAGTATTTTCTTAATTAAGATGATTTAAATATGAATCTTATGAGTAA 2944
Db |||||
QY 2828 TTCTCAGGATAGTATTTTCTTAATTAAGATGATTTAAATATGAATCTTATGAGTAA 2884
Db |||||

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2004, 02:10:46 ; Search time 1155 Seconds

(without alignments)

11637.488 Million cell updates/sec

Title: US-09-744-314-14

Perfect score: 3164

Sequence: 1 aggtttaaccatgaagaag.....acttgatttaaaaaaaaaa 3164

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124093041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002s.*

7: Geneseqn2003as.*

8: Geneseqn2003bs.*

9: Geneseqn2003cs.*

10: Geneseqn2004s.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3164	100.0	3164	3 AAZ58979	Aaz58979 Human cyt
2	2860.4	91.0	3108	7 ABQ77437	Abq77437 Human CGD
3	2567.8	81.2	3681	5 AAS69467	Aas69467 DNA encod
4	2515.2	79.5	2928	7 ABQ77424	Abq77424 Human CGD
5	2501.4	79.1	3915	5 AAS84445	Aas84445 DNA encod
6	2483.4	78.5	3296	7 ACD19410	Ac119410 cDNA enco
7	2414.6	76.3	2814	6 ABL67999	Ab167999 Ovary can
8	2414.6	76.3	2814	6 ABL68239	Ab168239 Human PRO
9	2414.6	76.3	2814	6 ABL96922	Ab196922 Gene #342
10	2414.6	76.3	2814	6 ABL95728	Ab195728 Human ang
11	2414.6	76.3	2814	9 ADD10624	Add10624 Human sec
12	2414.6	76.3	2814	9 ADD11584	Add11584 Human sec
13	2414.6	76.3	2814	9 ADD37377	Add37377 Human sec
14	2414.6	76.3	2814	10 ADE41585	Ade41585 Human sec
15	2397.2	75.8	2602	7 ACD13199	Acd13199 cDNA enco
16	2384.6	75.4	2762	5 AAS69466	Aas69466 DNA encod
17	2336.2	73.5	4345	5 AAF61716	Aaf61716 Human CFA
18	2303.4	72.8	2614	7 ACD13200	Acd13200 cDNA enco
19	2292.2	72.4	2775	5 AAS84444	Aas84444 DNA encod
20	2159.8	68.3	2351	4 AAI24168	Aai24168 Probe #14
21	2159.8	68.3	2351	4 ABA69288	Aba69288 Human foe
22	2159.8	68.3	2351	4 AAI49455	Aai49455 Probe #18
23	2159.8	68.3	2351	4 ABA51280	Aba51280 Human bre

24	2159.8	68.3	2351	4 ABA36214	Aba36214 Probe #14
25	2159.8	68.3	2351	4 AAK17575	Aak17575 Human bra
26	2159.8	68.3	2351	4 ABA3008	Abs43008 Human liv
27	2159.8	68.3	2351	5 AAI09732	Aai09732 Probe #97
28	2159.8	68.3	2351	6 ABA17476	Abs17476 Human gen
29	2148.6	67.9	2334	4 AAI24172	Aai24172 Probe #14
30	2148.6	67.9	2334	4 ABA69293	Aba69293 Human foe
31	2148.6	67.9	2334	4 AAI49459	Aai49459 Human bre
32	2148.6	67.9	2334	4 ABA51285	Aba51285 Human bre
33	2148.6	67.9	2334	4 ABA53012	Abs43012 Human liv
34	2148.6	67.9	2334	6 ABA17482	Abs17482 Human gen
35	2148.2	67.9	3263	5 AAS84443	Aas84443 DNA encod
36	2111.4	66.7	2588	7 AAT16043	Abt16043 NOXV rela
37	2026.2	64.0	2483	7 ABT16044	Abt16044 NOXV rela
38	1954.6	61.8	4292	2 AAX80481	Aax80481 Human sec
39	1954.6	61.8	4292	2 AAS92227	Aas92227 Human CDN
40	1954.6	61.8	4292	6 ABA90896	Aba90896 Human pol
41	1838.8	58.1	1986	4 AAI14980	Aai14980 Probe #49
42	1838.8	58.1	1986	4 ABA56714	Aba56714 Human foe
43	1838.8	58.1	1986	4 AAI36326	Aai36326 Probe #50
44	1838.8	58.1	1986	4 ABA46168	Aba46168 Human bre
45	1838.8	58.1	1986	4 ABA26329	Aba26329 Probe #47

ALIGNMENTS

RESULT 1

AAZ58979

ID AAZ58979 standard; cDNA; 3164 BP.

XX AC AAZ58979;

XX DT 08-MAY-2000 (first entry)

XX DE Human cytoskeletal protein (HCYT) encoding cDNA (clone 3768043).

XX KW Human cytoskeletal protein; HCYT; cell proliferation; immunological;

XX KW reproductive; nervous disorder; cytostatic; immunosuppressive; anti-HIV;

XX KW anti-diabetic; antiarteriosclerotic; dermatological; anti-inflammatory;

XX KW anti-infertility; vasotropic; cardiant; ss.

XX OS Homo sapiens.

XX FH Key

XX FT CDS

XX FT Location/Qualifiers

XX FT 352..2772

XX FT /*tag= a

XX PN WO200006730-A2.

XX PD 10-FEB-2000.

XX PF 30-JUL-1999; 99WO-US017167.

XX PR 31-JUL-1998; 98US-0155185P.

XX PR 04-AUG-1998; 98US-0160081P.

XX PR 19-AUG-1998; 98US-0155228P.

XX PA (INCY-) INCYTE PHARM INC.

XX PI Bandman O, Tang YT, Yue H, Corley NC, Guegler KJ, Azimzai Y;

XX PI Patterson C, Lal P, Baughn MR;

XX DR MPI; 2000-195297/17.

XX DR P-PSDB; AAY77574.

XX PT Human cytoskeletal proteins useful for diagnosing, treating preventing

XX PT cell proliferative, immunological, reproductive, developmental and

XX PT nervous disorders.

XX PS Claim 9; Page 79-80; 84pp; English.

XX CC The invention provides human cytoskeletal proteins (HCYT) and nucleic

acids encoding the proteins. The HCVt polypeptides can be expressed by standard recombinant methodology. The HCVt polypeptides, modulators and antibodies are useful for treating or preventing a disorder associated with decreased and increased expression or activity of HCVt in mammals. The polypeptides are also useful for diagnosing HCVt activity disorders such as cell proliferative, immunological, reproductive, developmental and nervous disorders. Sequences AA25874-981 represent cDNA clones encoding the HCVt polypeptides

841	DB	GATGAATTGATGGCGGATATTTC	AAAGACTATACAGTCTCTAGAGATGAGTCAAAATCC	900
901	QY	CTCAAAATCAAAATAGCTGAAGCC	AAATATCTCTGCAAGACATTTTAAATAGAGTGAAGAA	960
901	DB	CTCAAAATCAAAATAGCTGAAGCC	AAATATCTCTGCAAGACATTTTAAATAGAGTGAAGAA	960
961	QY	CGAGGGCTATAGCAATAAAGATGCTTT	GAATGAAAATTCCTCAACTTCAGACAAGGCCAT	1020
961	DB	CGAGGGCTATAGCAATAAAGATGCTTT	GAATGAAAATTCCTCAACTTCAGACAAGGCCAT	1020
1021	QY	AAACAGCTTTTTCAGCAAGAGCTGAAGTAT	TGGAAAGAGAGAGTGTGAGTGAACTTTATATAA	1080
1021	DB	AAACAGCTTTTTCAGCAAGAGCTGAAGTAT	TGGAAAGAGAGAGTGTGAGTGAACTTTATATAA	1080
1081	QY	CAGAAAATAACATTTGAAGACTCCAAAGTAC	ACGCGAGAAACAAAGTCTCGAATGATAAAGAA	1140
1081	DB	CAGAAAATAACATTTGAAGACTCCAAAGTAC	ACGCGAGAGAACAGTCTCGAATGATAAAGAA	1140
1141	QY	AATCACATCAAGACCTGTAGTGGACATCTTG	CCAAATGATGAAGAGATCAGCGTGTGTGCTT	1200
1141	DB	AATCACATCAAGACCTGTAGTGGACATCTTG	CCAAATGATGAAGAGATCAGCGTGTGTGCTT	1200
1201	QY	GAGAAAGACAAACGGATGATTAACCTGGAA	TTAGAAGTGAAACAGTGAATCGGAAAT	1260
1201	DB	GAGAAAGACAAACGGATGATTAACCTGGAA	TTAGAAGTGAAACAGTGAATCGGAAAT	1260
1261	QY	GGTGTACTTAGATTAATCCTCGAAAGAGAGCT	TTGAAGAAACTGATTCATGCTGCTAAG	1320
1261	DB	GGTGTACTTAGATTAATCCTCGAAAGAGAGCT	TTGAAGAAACTGATTCATGCTGCTAAG	1320
1321	QY	TTAAATGCTCTTTTAAAAACCTTTAGAAGG	AGAAACCAATTTATATTCAAGTTGTCT	1380
1321	DB	TTAAATGCTCTTTTAAAAACCTTTAGAAGG	AGAAACCAATTTATATTCAAGTTGTCT	1380
1381	QY	GAAATTGATAAAACAAAGAGAGCTTACAG	AGCATTTAAAAATCTTCAGACTCAACA	1440
1381	DB	GAAATTGATAAAACAAAGAGAGCTTACAG	AGCATTTAAAAATCTTCAGACTCAACA	1440
1441	QY	GCATCTTTCAGTCAAGAAACACACATTTT	TGAAAAATGAGAATCAGAAGCTTCAACAGAAA	1500
1441	DB	GCATCTTTCAGTCAAGAAACACACATTTT	TGAAAAATGAGAATCAGAAGCTTCAACAGAAA	1500
1501	QY	CTTAAAGTAAATGACTGAATATATCAAG	AAATGAAATGAAATCCACAGGAAATTAACA	1560
1501	DB	CTTAAAGTAAATGACTGAATATATCAAG	AAATGAAATGAAATCCACAGGAAATTAACA	1560
1561	QY	GTAGAGGAAAAATTCGGTTAGAGAAAG	AGAGAACTTTCTAAAGTAGATGAAAAAGATC	1620
1561	DB	GTAGAGGAAAAATTCGGTTAGAGAAAG	AGAGAACTTTCTAAAGTAGATGAAAAAGATC	1620
1621	QY	AGCCATGCCACTGAAGAGCTGGAGACCTAT	TAGAAACGAGCCAAAGATCTTCAGAGAGAA	1680
1621	DB	AGCCATGCCACTGAAGAGCTGGAGACCTAT	TAGAAACGAGCCAAAGATCTTCAGAGAGAA	1680
1681	QY	TTGGAGAGAACTATTCAATCTTTTCAAG	GGCAGATTTTCCCATGAGAAAAAGCACAT	1740
1681	DB	TTGGAGAGAACTATTCAATCTTTTCAAG	GGCAGATTTTCCCATGAGAAAAAGCACAT	1740
1741	QY	GATAATTGGTTGGCAGCTCGGAATCTG	AAAGAAACCTCAATGATTTAAGGAAGAGAAAT	1800
1741	DB	GATAATTGGTTGGCAGCTCGGAATCTG	AAAGAAACCTCAATGATTTAAGGAAGAGAAAT	1800
1801	QY	GCTCAACAACAGACAAAAATTAAC	TGAAAACAGAGCTTTTAAATTTTGAATTTT	1860
1801	DB	GCTCAACAACAGACAAAAATTAAC	TGAAAACAGAGCTTTTAAATTTTGAATTTT	1860
1861	QY	CCTTATGCACTCGATGTTCCAAATACAG	ATTTGGCAGAGAGCATTTCCCATATGTCCTC	1920
1861	DB	CCTTATGCACTCGATGTTCCAAATACAG	ATTTGGCAGAGAGCATTTCCCATATGTCCTC	1920
1921	QY	TCACCAATTGGGTTGGCTTTCATCTG	AAAAACAGAGCTTTTCTCTCTCTCTCTCTCT	1980
1921	DB	TCACCAATTGGGTTGGCTTTCATCTG	AAAAACAGAGCTTTTCTCTCTCTCTCTCTCT	1980

QY	1981	GAGGGTCCACTCAGACTCTCACTCTTGCTTCCAGGGGGAGGAGGAGGCTCAGAGGC	2040
Db	1981	GAGGGTCCACTCAGACTCTCACTCTTGCTTCCAGGGGGAGGAGGAGGCTCAGAGGC	2040
QY	2041	CCAGGGAATCCTCTGCAGCACTCAGATTACCAATGAAAGAGGAGAAATCAAGCTGTATAGG	2100
Db	2041	CCAGGGAATCCTCTGCAGCACTCAGATTACCAATGAAAGAGGAGAAATCAAGCTGTATAGG	2100
QY	2101	TTAAACCGATCCTCATAGGGGTCCTCTGCACACTGGGTCTCTCTCACCTCCATGGGACCAG	2160
Db	2101	TTAAACCGATCCTCATAGGGGTCCTCTGCACACTGGGTCTCTCTCACCTCCATGGGACCAG	2160
QY	2161	GACCGTAGGATGATGTTTCTCGCCAGGACAAATCATATCCTGTATCAGCCCTTCTCTCCA	2220
Db	2161	GACCGTAGGATGATGTTTCTCGCCAGGACAAATCATATCCTGTATCAGCCCTTCTCTCCA	2220
QY	2221	CAAAAGCAAGACAGATTTTGTCTTAATTTCTGGTAGACTCTCTGGACACGAGAATCTAGA	2280
Db	2221	CAAAAGCAAGACAGATTTTGTCTTAATTTCTGGTAGACTCTCTGGACACGAGAATCTAGA	2280
QY	2281	AGTTTTTAATGCTTCTTTGGATAAAATGGATGGGTCAATGCTTCAGAAATGGAATCC	2340
Db	2281	AGTTTTTAATGCTTCTTTGGATAAAATGGATGGGTCAATGCTTCAGAAATGGAATCC	2340
QY	2341	AGTAGAAATGATACCAAAGATGATCTTGGTAAATTTAAATGTGCTGATTCATCTCTCCCT	2400
Db	2341	AGTAGAAATGATACCAAAGATGATCTTGGTAAATTTAAATGTGCTGATTCATCTCTCCCT	2400
QY	2401	GCTGAAAATGAAGGCACCTGGCCCTGGCTTTGTTTCTCCACCTCTTGCTCCAATCAGAGGT	2460
Db	2401	GCTGAAAATGAAGGCACCTGGCCCTGGCTTTGTTTCTCCACCTCTTGCTCCAATCAGAGGT	2460
QY	2461	CCATTGTTTCCAGTGGATGCAAGAGGCCCATTTCTTGAGAAAGAGGACCTCTTTCCCCCCA	2520
Db	2461	CCATTGTTTCCAGTGGATGCAAGAGGCCCATTTCTTGAGAAAGAGGACCTCTTTCCCCCCA	2520
QY	2521	CTCTCTCCAGGAGCCATGTTTGGAGCTTCTCGAGATTAATTTTCCACGAGGATTTCCCA	2580
Db	2521	CTCTCTCCAGGAGCCATGTTTGGAGCTTCTCGAGATTAATTTTCCACGAGGATTTCCCA	2580
QY	2581	GGTCCACACCTGCTCCATTGCAATGAGAAATGTCTATCCACGAGGGGTGTTCTCTCCT	2640
Db	2581	GGTCCACACCTGCTCCATTGCAATGAGAAATGTCTATCCACGAGGGGTGTTCTCTCCT	2640
QY	2641	TACCTTCCCGCAGACCTCGATTTTTCCTCCCGCACCCCCCAATTCGAGAGTGAAATGAG	2700
Db	2641	TACCTTCCCGCAGACCTCGATTTTTCCTCCCGCACCCCCCAATTCGAGAGTGAAATGAG	2700
QY	2701	TTCCCTCTCAGTTTGATTTCCACCTTCAAAATGAGGCTGCTACTGAAATCCAGAAACACAG	2760
Db	2701	TTCCCTCTCAGTTTGATTTCCACCTTCAAAATGAGGCTGCTACTGAAATCCAGAAACACAG	2760
QY	2761	CAAGAAACCTGACAAATATTTTGTCTCTTCCAAAAGTAAATTTTGA CTGATCTCATTTTCA	2820
Db	2761	CAAGAAACCTGACAAATATTTTGTCTCTTCCAAAAGTAAATTTTGA CTGATCTCATTTTCA	2820
QY	2821	GTTTAAAGTAACTGCTGTACTTAAAGTAACTACCTTTTGTCTCAAAATGGAAGCTTAAATGA	2880
Db	2821	GTTTAAAGTAACTGCTGTACTTAAAGTAACTACCTTTTGTCTCAAAATGGAAGCTTAAATGA	2880
QY	2881	ATTATAATTTCTCAGGATAGTATTTTGTAAATFAAAGATGATTTAAATATGAATCTTATGAG	2940
Db	2881	ATTATAATTTCTCAGGATAGTATTTTGTAAATFAAAGATGATTTAAATATGAATCTTATGAG	2940
QY	2941	TAAATTAATTCAAATTTTATTTTATAGCCGTATACTATTTTCAATTTGATTTAATCCACTATT	3000
Db	2941	TAAATTAATTCAAATTTTATTTTATAGCCGTATACTATTTTCAATTTGATTTAATCCACTATT	3000
QY	3001	ATATAAACAAATAGTGGGAGTTTTTATATATGTAATCTTTTCAGGTGGGAGGCTTTAAATTC	3060
Db	3001	ATATAAACAAATAGTGGGAGTTTTTATATGTAATCTTTTCAGGTGGGAGGCTTTAAATTC	3060

Qy	3061	TGAAGTCTGTGCTCTTATGCCAAGAACTGATTTACTGTGTTGTCGACAAATGTGAAG	3122
Db	3061	TGAAGTCTGTGCTCTTATGCCAAGAACTGATTTACTGTGTTGTCGACAAATGTGAAG	3120
Qy	3121	TAACTTTATGCTTAAATAAATATATAGTTGATTTAAAAA	3164
Db	3121	TAACTTTATGCTTAAATAAATATATAGTTGATTTAAAAA	3164
RESULT 2			
ABQ77437			
ID	ABQ77437	standard; cDNA; 3108 BP.	
XX	ABQ77437;		
XX			
XX			
DT	10-MAY-2003	(first entry)	
XX			
DE	Human CGDB	cDNA 7506167CB1 SEQ ID 51.	
XX			
KW	Human; cell growth, differentiation and death; CGDD; cardiact; cancer;		
KW	cytostatic; neuroprotective; gene therapy; diagnosis; cardiovascular;		
KW	neurological disorder; gene; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	W02003014322-A2.		
XX			
PD	20-FEB-2003.		
XX			
PF	08-AUG-2002; 2002WO-US025465.		
XX			
PR	08-AUG-2001; 2001US-0311017P.		
PR	17-AUG-2001; 2001US-0313070P.		
PR	17-AUG-2001; 2001US-0313071P.		
PR	24-AUG-2001; 2001US-0314678P.		
PR	31-AUG-2001; 2001US-0316692P.		
PR	07-SEP-2001; 2001US-0317913P.		
PR	14-SEP-2001; 2001US-0322182P.		
PR	07-DEC-2001; 2001US-0340747P.		
PR	20-DEC-2001; 2001US-0342761P.		
PR	29-MAR-2002; 2002US-0369129P.		
XX			
FA	(INCY-) INCYTE GENOMICS INC.		
XX			
PI	Azimzal Y, Barroso I, Baughn MR, Becha SD, Borowsky ML;		
PI	Duggan BM, Elliott VS, Emerling BM, Forsythe IU, Gietzen KU;		
PI	Govad AE, Graul FC, Griffin JA, Gururajan R, Hafalia AJA, Ison CH;		
PI	Xable AE, Khan PA, Lee S, Lee SY, Li JX, Reddy R, Richardson TW;		
PI	Sprague WW, Swarnakar A, Tang YT, Warren BA, Xu Y, Yao MG, Yue H;		
PI	Yue H;		
XX			
XX	WPI; 2003-268197/26.		
DR	P-PSDB; ABG74698.		
XX			
PT	New polypeptide associated with cell growth, differentiation and death,		
PT	useful for preparing a composition for diagnosing or treating a disease		
PT	e.g., cardiovascular or neurological disorders or cancer.		
XX			
PS	Claim 106; Page 264-265; 267pp; English.		
XX			
CC	This invention describes a novel polypeptide associated with cell growth,		
CC	differentiation and death (CGDD) which has cardiant, cytostatic and		
CC	neuroprotective activity. The polypeptides of the invention are useful		
CC	for preparing a composition for diagnosing, or treating a disease or		
CC	condition associated with decreased expression or overexpression of		
CC	functional CGDD e.g., cardiovascular or neurological disorders or cancer.		
CC	The polynucleotides of the invention can be used for gene therapy.		
CC	ABQ77414-ABQ77441 encode the human CGDD polypeptides represented in		
CC	ABG74675-ABG74701		

X
Q Sequence 3108 BP; 1011 A; 607 C; 673 G; 817 T; 0 U; 0 Other;
Query Match 91.0%; Score 2880.4; DB 7; Length 3108;

Best Local Similarity 95.7%; Pred. No. 0;		Matches 3023; Conservative 0; Mismatches 6; Indels 129; Gaps 1;	
QY	1 AGGTTTAAATCCATGAAAGAACACAGCAATTTTAAAAGTGTATTTCACCAAAAAATAAAGCTT 60		
Db	72 AGGTTTAAATCCATGAAAGAACACAGCAATTTTAAAAGTGTATTTCACCAAAAAATAAAGCTT 131		
QY	61 CAAAAATATGTGATGTGAAAAATCTCCAGAATTAAGCGGGCCGGGCTCAGACAGCGCTGC 120		
Db	132 CAAAAATATGTGATGTGAAAAATCTCCAGAATTAAGCGGGCCGGGCTCAGACAGCGCTGC 191		
QY	121 CTCAGGATGTAAAGTGTAAACAAGAGGCCAGGGAGGTGTGGGGGACAACATGGGCCCTG 180		
Db	192 CTCAGGATGTAAAGTGTAAACAAGAGGCCAGGGAGGTGTGGGGGACAACATGGGCCCTG 251		
QY	181 TGAAGGCTGTGGGTGCGCGGTTCCACAGTCCCGCGCAGCCCGCTCCACAGTGTGCTCG 240		
Db	252 TGAAGGCTGTGGGTGCGCGGTTCCACAGTCCCGCGCAGCCCGCTCCACAGTGTGCTCG 311		
QY	241 CTCGGGTGGTGTGTCACGTGCGCATTCGGGTTCAGACCCCAAGGCTCGGTGTTCTCCACC 300		
Db	312 CTCGGGTGGTGTGTCACGTGCGCATTCGGGTTCAGACCCCAAGGCTCGGTGTTCTCCACC 371		
QY	301 GCTTGTGTGGCAGATGTTACTCGGTGACCGCCAGAGCAGCCTCGACGCTATGAGAGAG 360		
Db	372 GCTTGTGTGGCAGATGTTACTCGGTGACCGCCAGAGCAGCCTCGACGCTATGAGAGAG 431		
QY	361 CTTGGTGTACTACCCCTCAGCCCTACCTCGGGCTGGTCTCTGGAGGAGCTAGCGCAGATGTGTG 420		
Db	432 CTTGGTGTACTACCCCTCAGCCCTACCTCGGGCTGGTCTCTGGAGGAGCTAGCGCAGATGTGTG 491		
QY	421 GCAGCACTACTGAGAGTATGAGACCCAGATGAGAAATCCTTATGGTTTTTCATCGGAACCTG 480		
Db	492 GCAGCACTACTGAGAGTATGAGACCCAGATGAGAAATCCTTATGGTTTTTCATCGGAACCTG 551		
QY	481 GTGGTATGTCAGCTGTTATTTGGATTTTTTGTGTCTCTCTTTTTTGTGGAGAGTTT 540		
Db	552 GTGGTATGTCAGCTGTTATTTGGATTTTTTGTGTCTCTCTTTTTTGTGGAGAGTTT 611		
QY	541 AGATCGGTTAGGAGTCGGCTTTACGTGGGAGAGAGCAAAAACCTGGTGCAACGCTTTCT 600		
Db	612 AGATCGGTTAGGAGTCGGCTTTACGTGGGAGAGAGCAAAAACCTGGTGCAACGCTTTCT 671		
QY	601 GGACTAATTGAAGAAAAATGTAAACTACTTTGAAAAATTTAGCCCTTATTCAAAAAGAGTAT 660		
Db	672 GGACTAATTGAAGAAAAATGTAAACTACTTTGAAAAATTTAGCCCTTATTCAAAAAGAGTAT 731		
QY	661 GAAGGCTATGAAGTAGAGTCATCTTTTAGAGATGCCAGCTTTGAGAGGAGGCGAGAGAA 720		
Db	732 GAAGGCTATGAAGTAGAGTCATCTTTTAGAGATGCCAGCTTTGAGAGGAGGCGAGAGAA 791		
QY	721 GAAGCACAAGTTTGGAGGCAACCTGTGAAAGCTGAACAGGTCCTAATTCCTGAACCTTGAG 780		
Db	792 GAAGCACAAGTTTGGAGGCAACCTGTGAAAGCTGAACAGGTCCTAATTCCTGAACCTTGAG 851		
QY	781 GATGAAATCCTCTCTCTAGAAAAAGACTTAAAAACAAGAGAAATCTAAACATTTCTCAACAA 840		
Db	852 GATGAAATCCTCTCTCTAGAAAAAGACTTAAAAACAAGAGAAATCTAAACATTTCTCAACAA 911		
QY	841 GATGAAATGTCGGGATATTTCAAAGATATACAGTCTCTAGAGATGAGTCAAATCC 900		
Db	912 GATGAAATGTCGGGATATTTCAAAGATATACAGTCTCTAGAGATGAGTCAAATCC 971		
QY	901 CTCAAATCACAAATAGCTGAAAGCCAAATCATCTTGCAGAGCATTTAAAAATGAGTGAAGAA 960		
Db	972 CTCAAATCACAAATAGCTGAAAGCCAAATCATCTTGCAGAGCATTTAAAAATGAGTGAAGAA 1031		
QY	961 CGACGGGCTATAGCAATTAAGATGCTTTGAAATGAAATTTCTCAACTTCAGACAGCCAT 1020		
Db	1032 CGACGGGCTATAGCAATTAAGATGCTTTGAAATGAAATTTCTCAACTTCAGACAGCCAT 1091		
QY	1021 AAACAGCTTTTTTCAGCAAGAGAGCTGAAGTATGAAAAAGAGAGTGAAGTGAACCTTAATAAA 1080		

1092	DB	AAACAGCTTTTTCACGAGAAAGCTGAAAGTATGGAAGAGAGAGTGAAGTGAACCTTAATAAA	1155
1081	QY	CAGAAAATAACATTTGAAGACTCCAAAGTACACGAGAAACAAGTTCTGAAATGATAAGAA	1140
1152	DB	CAGAAAATAACATTTGAAGACTCCAAAGTACACGAGAAACAAGTTCTGAAATGATAAGAA	1211
1141	QY	AATCATCAAGACCTGACTGGACACTTGGCAATGATGAAGATCAGGCTGCTGCTT	1200
1212	DB	AATCATCAAGACCTGACTGGACACTTGGCAATGATGAAGATCAGGCTGCTGCTT	1271
1201	QY	GAGGAAGACACAACGGATGATGATAACCTGGAAATTAGAAGTGAACAGTGAATCGGAAT	1260
1272	DB	GAGGAAGACACAACGGATGATGATAACCTGGAAATTAGAAGTGAACAGTGAATCGGAAT	1331
1261	QY	GGTGCTTACTTAGATAAATCCTCCAAAAGAGCTTTGAAGAACTGATTCATGCTGCTAAG	1320
1332	DB	GGTGCTTACTTAGATAAATCCTCCAAAAGAGCTTTGAAGAACTGATTCATGCTGCTAAG	1391
1321	QY	TTAAATGCTCTTTTAAACCTTTAGAAGAGAGAAAGAAACCAATTTTATTCAGTTGTCT	1380
1392	DB	TTAAATGCTCTTTTAAACCTTTAGAAGAGAGAAAGAAACCAATTTTATTCAGTTGTCT	1451
1381	QY	GAGTTGATAAAACAAGGAAGAGCTTACAGAGCATTTAAAAATCTTCAGACTCAACAA	1440
1452	DB	GAGTTGATAAAACAAGGAAGAGCTTACAGAGCATTTAAAAATCTTCAGACTCAACAA	1511
1441	QY	GCATCTTTCAGTCAAGAAACACACATTTTGAAATGAGAAATCAGAGCTTCAACAGAA	1500
1512	DB	GCATCTTTCAGTCAAGAAACACACATTTTGAAATGAGAAATCAGAGCTTCAACAGAA	1571
1501	QY	CTTAAAGTAATGACTGAAATTTATATCAAGAAATGAAATGAAATCTCCACAGGAAATTAA	1560
1572	DB	CTTAAAGTAATGACTGAAATTTATATCAAGAAATGAAATGAAATCTCCACAGGAAATTAA	1631
1561	QY	GTAGAGGAAATTTATCGTTTAGAAGAGAGAAACTTCTAAAGTAGATGAAGATC	1620
1632	DB	GTAGAGGAAATTTATCGTTTAGAAGAGAGAAACTTCTAAAGTAGATGAAGATC	1691
1621	QY	AGCCATGCCACTGAAGAGCTGGAGACCTATAGAAAGCGAGCCAAAGATCTTGAAGAGAA	1680
1692	DB	AGCCATGCCACTGAAGAGCTGGAGACCTATAGAAAGCGAGCCAAAGATCTTGAAGAGAA	1751
1681	QY	TTGGAGAGACTATTCACTCTTATCAGGCGCAGATTTTCCCATGAGAAAGACAT	1740
1752	DB	TTGGAGAGACTATTCACTCTTATCAGGCGCAGATTTTCCCATGAGAAAGACAT	1811
1741	QY	GATAATTTGGTTGGCAGCTCGAAATGCTGAAAGAAACCTCAATGATTTAAGGAAGAAAT	1800
1812	DB	GATAATTTGGTTGGCAGCTCGAAATGCTGAAAGAAACCTCAATGATTTAAGGAAGAAAT	1871
1801	QY	GCTTCACACAGACAAAATTAACCTGAACAGAGCTTAAATTTGAATTTAGGAAGAGAT	1860
1872	DB	GCTTCACACAGACAAAATTAACCTGAACAGAGCTTAAATTTGAATTTAGGAAGAGAT	1931
1861	QY	CCTTATGCACTCGATGTTCCAAATACAGCAATTTGGCAGAGAGCATTTCCCATATGTFCCC	1920
1932	DB	CCTTATGCACTCGATGTTCCAAATACAGCAATTTGGCAGAGAGCATTTCCCATATGTFCCC	1957
1921	QY	TCACCATTGGTTGGCCTTCATCTGAACAAGAGCTTTCTCTCTCTCCCACTTTCTTG	1980
1968	DB	-----	1967
1981	QY	GAGGGTCCACTCAGACTCTCACCTTTGCTTCCAGGGGAGGAGAGAGGCTTCAAGAGC	2040
1968	DB	-----AGAGGCTCACGAGC	1982
2041	QY	CCAGGGAATCCTCTGGACCATCAGATTTACCAATGAAAGAGAGAGATCAAGCTGTGATAGG	2100
1983	DB	CCAGGGAATCCTCTGGACCATCAGATTTACCAATGAAAGAGAGAGATCAAGCTGTGATAGG	2042
2101	QY	TTAACCGATCCTCATAGGGTCCCTCTGACACTGGGTCTCTGTCACTCCATCGGACAG	2160
2043	DB	TTAACCGATCCTCATAGGGTCCCTCTGACACTGGGTCTCTGTCACTCCATCGGACAG	2102

2161	QY	GACCGTAGGATGATGTTTCTCCGCGCAGGACAATCATATCTCTGATTCAGCCCTTCCTCCA	2220
2103	Db	GACCGTAGGATGATGTTTCTCCGCGCAGGACAATCATATCTCTGATTCAGCCCTTCCTCCA	2162
2221	QY	CAAGGCGAAGACAGATTTGTTCTAAATCTCGTAGACTGTCTGGACACGACGAACTCAGA	2280
2163	Db	CAAGGCGAAGACAGATTTGTTCTAAATCTCGTAGACTGTCTGGACACGACGAACTCAGA	2222
2281	QY	AGTTTTAATATGCCCTCTTTGGATAAAATGGATGGGTCAATGCCCTTCAGAAATGGAATCC	2340
2223	Db	AGTTTTAATATGCCCTCTTTGGATAAAATGGATGGGTCAATGCCCTTCAGAAATGGAATCC	2282
2341	QY	AGTAGAATGATACCAAGATGATCTTGGTAATTTAAATGTGCGCTGATTCATCTCCCT	2400
2283	Db	AGTAGAATGATACCAAGATGATCTTGGTAATTTAAATGTGCGCTGATTCATCTCTCCCT	2342
2401	QY	GCTGAAAATGAAGCCACTGGCCCTGGCTTGTCTCTCCACCTCTTGTCTCCAATCAGAGGT	2460
2343	Db	GCTGAAAATGAAGCCACTGGCCCTGGCTTGTCTCTCCAATCAGAGGT	2402
2461	QY	CGATGTTTTCCAGTGGATGCAAGAGGCCCATCTTTGAGAAGAGGACCTCTCTTTCCGCCCA	2520
2403	Db	CGATGTTTTCCAGTGGATGCAAGAGGCCCATCTTTGAGAAGAGGACCTCTCTTTCCGCCCA	2462
2521	QY	CCTCCTCCAGGAGCCATGTTTGGAGCTCTCGAGATTATTTTCCACCAAGGATTTCCCA	2580
2463	Db	CCTCCTCCAGGAGCCATGTTTGGAGCTCTCGAGATTATTTTCCACCAAGGATTTCCCA	2522
2581	QY	GGTCCACCACTGCTCCATTTGCAATGAGAAATGCTATCCACCGAGGGTTTTCTCTCT	2640
2523	Db	GGTCCACCACTGCTCCATTTGCAATGAGAAATGCTATCCACCGAGGGTTTTCTCTCT	2582
2641	QY	TACCTTCCCGCAAGACCTGGATTTTCCCGCCACCCACATCTCGAAGGTAGAGTGAG	2700
2583	Db	TACCTTCCCGCAAGACCTGGATTTTCCCGCCACCCACATCTCGAAGGTAGAGTGAG	2642
2701	QY	TTCCCTCAGGTTTGATTCACCTTCAAAATGAGCCTGCTACTGAAATCCAGAACACAG	2760
2643	Db	TTCCCTCAGGTTTGATTCACCTTCAAAATGAGCCTGCTACTGAAATCCAGAACACAG	2702
2761	QY	CAAGAAACCTGACAATATTTTTCCTCTCTTCAAAAGTAATTTTGACTGATCTCATTTTCA	2820
2703	Db	CAAGAAACCTGACAATATTTTTCCTCTCTTCAAAAGTAATTTTGACTGATCTCATTTTCA	2762
2821	QY	GTTTAAAGTAACCTGTTTACTTAAAGTGATTACACTTTTGTCTCAAAATGAACTTAATGGA	2880
2763	Db	GTTTAAAGTAACCTGTTTACTTAAAGTGATTACACTTTTGTCTCAAAATGAACTTAATGGA	2822
2881	QY	ATTATTAATCTCAGGATAGPATTTTGTGAAATAAAGATGATTTAAATATGAACTTATGAG	2940
2823	Db	ATTATTAATCTCAGGATAGPATTTTGTGAAATAAAGATGATTTAAATATGAACTTATGAG	2882
2941	QY	TAAATTTATTCAAATTTATTTTGAACGGTATACTATTTCAAATTTGATTAATCCACTATT	3000
2883	Db	TAAATTTATTCAAATTTATTTTGAACGGTATACTATTTCAAATTTGATTAATCCACTATT	2942
3001	QY	ATATAACAATAGTGGAGTTTTATATATGTAATCTTTTCAGGTGGGAGGCTTTAAATTC	3060
2943	Db	ATATAACAATAGTGGAGTTTTATATATGTAATCTTTTCAGGTGGGAGGCTTTAAATTC	3002
3061	QY	TGAAGTCTGTCTTTATGCGAAGACTGATTTTACTGTGTTGTGGCAAAATGTCGAAG	3120
3003	Db	TGAAGTCTGTCTTTATGCGAAGACTGATTTTACTGTGTTGTGGCAAAATGTCGAAG	3062
3121	QY	TAACTTTATGCTTAAATAAAATTAAGTTGATTTTAAAAA	3158
3063	Db	TAACTTTATGCTTAAATAAAATTAAGTTGATTTTAAAAA	3100

RESULT 3

AAS69467

ID AAS69467 standard; cdna; 3691 BP.

RESULT 3
AAS69467
ID AAS

XX	AAS69467;
AC	
XX	13-FEB-2002 (first entry)
DT	
XX	DNA encoding novel human diagnostic protein #5271.
XX	
XX	Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW	food supplement; medical imaging; diagnostic; genetic disorder; ss.
KW	
OS	Homo sapiens.
OS	
PN	WO200175067-A2.
PN	
PD	11-OCT-2001.
XX	
XX	30-MAR-2001; 2001WO-US008631.
PF	
XX	31-MAR-2000; 2000US-00540217.
PR	
PR	23-AUG-2000; 2000US-00649167.
PA	(HYSE-) HYSEQ INC.
XX	
XX	Drmanac RT, Liu C, Tang YT;
PI	
XX	WPI; 2001-639362/73.
DR	P-PSDB; ABG05280.
DR	
XX	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensic, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity.
XX	
XX	Claim 1; SEQ ID NO 5271; 103pp; English.
PS	
CC	The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC	sequences. (I) is useful as hybridisation probes, polymerase chain
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC	and in recombinant production of (II). The polynucleotides are also used
CC	in diagnostics as expressed sequence tags for identifying expressed
CC	genes. (I) is useful in gene therapy techniques to restore normal
CC	activity of (II) or to treat disease states involving (II). (II) is
CC	useful for generating antibodies against it, detecting or quantitating a
CC	polypeptide in tissue, as molecular weight markers and as a food
CC	supplement. (II) and its binding partners are useful in medical imaging
CC	of sites expressing (II). (I) and (II) are useful for treating disorders
CC	involving aberrant protein expression or biological activity. The
CC	polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. AAG64197-AAS94564 represent novel human diagnostic
CC	coding sequences of the invention. Note: The sequence data for this
CC	patent did not appear in the printed specification, but was obtained in
CC	electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 3681 BP; 1134 A; 685 C; 805 G; 1056 T; 0 U; 1 Other;
	Query Match 81.2%; Score 2567.8; DB 5; Length 3681;
	Best Local Similarity 92.9%; Pred. No. 0;
	Matches 2881; Conservative 0; Mismatches 168; Indels 53; Gaps 16;
Qy	90 CTAAAGCGCGCGGGCTCAGACCAGCGCTCCCTCAGGATGTAAAGTGTAAACAAGAGGCC 149
Dd	57 CTTGTGGGTGGGTTCGGACCTGCCTCGGATGTAAAGTGTAAACAAGAGGTC 116
Qy	150 AGGGAGAGTGTTGGGGACAACTGGCGCTGTGAGGCGCTGTGGGTGCCCGCTTCCCCAG 209
Dd	117 GGGATGG-----CAGCGTAGGCTGTGAGGCGCTGGGGTGCCTGTCCCCAG 166
Qy	210 CTCCTCCCGGAGCCCGCTCACACTGTCGCTCGGTTGGTTGTCAAGTCGCGATTCCG 269

Db 167 CTCCCGCCGACCGCGCTCCCGAGTGGTCCACTCC-----GGTTCGCGGTGCGGATTGG 222
Qy 270 GTTCAGAGCCCAAGAGCTGGTCTCTCCACGCTGTTGT-----GGCCAGTGT 319
Db 223 GTTCGGAACCGAAGAGCTGTGTCTCTCGCGCTCAATTTGTGCGCCCGACAGCGCGGGTT 282
Qy 320 ACTGCGGTGACCGCCAGAGCAGCCTCGACGCTATCGAGAGCGCTGGTGTCTACCCCTCAGC 379
Db 283 ACTGTGGCCACACAGAGCAGCTTTGGCGCTATGGAGGAGCCCGGGCTACCCCTCAAC 342
Qy 380 CTACTCGGGGTGGTCTGGAGGAGCTACGAGAGTTGTGCGAGCACTACTCGAGAGTA 439
Db 343 CGTATTTGGGGTGTCTCTCGAGAGAGCTACGAGAGTTGTGGCAGCACTGCTGAGGTA 402
Qy 440 TGAGACAGATGAGAAATCCTTATGTTTCCATCGGAACCTGTGTGTATGTGCGAGCTGTTA 499
Db 403 TGAGACAGATCTTAATCTTTATGTTTTCCTATGGGAATTTGATGATATGTGCGAGCTGTTG 462
Qy 500 TTTGGATTTTGTGTCTCTCTTTTGTGAGAGTTTGTAGTCGGTATAGGAGTCGC 559
Db 463 TTTGGATTTTGTGTCTCTCTTTTGTGAGAGTTTGTAGTCGGTATAGGAGTCGC 522
Qy 560 TTTTACGTGGAAAGAGAGCAAAACCTTTGTGCAACGCTTTCTGAGCTAATTTGAAGAAAAT 619
Db 523 TTTATGTGGAGCAGAGAGAAAAGCTTGTCTTAATGCTTTCTGACTAATTTGAAGAAAAT 582
Qy 620 GTAACTACTGAAAATTTAGCCTTATTTCAAAAGAGTATGAGGCTATGAGTAGAGT 679
Db 583 GTAACTACTGAAAATTTAGCCTTATTTCAAAAGAGTATGAGGCTATGAGTAGAGT 642
Qy 680 CATCTTTAGAGTAGCCAGCTTTTGAAGAGGAGCAGAGAAAGCAGAACTTTGGAGG 739
Db 643 CATCTTTAAGGATGCCAGCTTTGAGAAGGAGGCAAC---AGAAGCACAAGTTTGGAGG 699
Qy 740 CAACTGTGAAAGCTGAACAGCTCAATTTCTGAAGTTGAGGATGAAATCTCTGCTAG 799
Db 700 CAACTGTGAAAGCTGAACAGCTCAATTTCTGAAGTTGAGGATGAAATCTCTGCTAG 759
Qy 800 AAAAAAGCTTTAAACAAAGAGAAATCTAAACATTTCTCAACAAGATGAATTTGATGGGGATA 859
Db 760 AAAAAAGCTTTAAACAAAGAGAAATCTAAACATTTCTCAACAAGATGAATTTGATGGGGATA 819
Qy 860 TTTCAAAAGTATACAGCTCTAGAGATGAGTCAAAATCCCTCAAAATCACAATAGCTG 919
Db 820 TTTCAAAAGGATACAGCTCTAGAGATGAGTCAAAATCCCTCAAAATCACAATAGCTG 879
Qy 920 AAGCAAAATCATCTGCAACACATTTAAATAGTCAAGAACGACGGGCTATAGCAATAA 979
Db 880 AAGCAAAATGACCTTCCAGATATTTCCATGAATGAGNACGATGAAGATAGCAATAA 939
Qy 980 AAGATGCTTTGAATGAAAATTTCACTTCAGACAGCCATTAACAGCTTTTTCAGCAAG 1039
Db 940 AAGATGCTTTGAATGAAAATTTCACTTCAGGAAAGCCAGAAACAGCTTTT---GCAAG 996
Qy 1040 AAGCTGAAGTATGAAAGGAGAGTGAAGTCAATTAATAACAGAAAATAACATTTGAAG 1099
Db 997 AAGCTGAAGTATGAAAGGAGAACAGTGAAGTGAATTTAATAACAGAAATGAACATTTGAAG 1056
Qy 1100 ACTCAAAGTACCGCAGAAACAAGTCTGAATGATTAAGAAAATCAATCAAGACCCCTGA 1159
Db 1057 ACTCAAAGTACATGCAAGAACAAAGTCTGAATGATTAAGAAAATCAATCAAGACCTGA 1116
Qy 1160 CTGAGACACTTGGCAATGATGAAGATCAGCTGCTGTGTTGAGAGACACAAAGGATG 1219
Db 1117 CTGAGACACTTGGCAATGATGAAGATGAGGCTGCTGTGTTGAGAGACACAAAGGATG 1176
Qy 1220 ATGATAACTTGGAAATTTAGAAAGTGAACAGTGAATCGGAAAATGTTGCTTACTTGAATAATC 1279
Db 1177 ATGATAACTTGGAAATTTAGAAATGAACAGTGAATCGGAAAATGTTGCTTACTTGAATAATC 1236
Qy 1280 CTCAAAAGAGCTTTGAAGAACTGATTCATGCTGCTAGTTAAATGCTTTTAAAAA 1339
Db 1237 CTCAAAAGAGGAGCTTTGAAGAACTGATTCATGCTGCTAGTTAAATGCTTTTAAAAA 1296

Qy 1340 CCTTAGAAGGAGAGAAACCAAAATTTTATTCAGTTGTCTGAAGTTGATATAAAACAAAGG 1399
Db 1297 CCTTAGAAGGAGAGAAACCAAAATTTTATTCAGTTGTCTGAAGTTGATATAAAACAAAGG 1356
Qy 1400 AAGAGCTTACAGAGCATATTTAAAAATCTTCAGACTCAACAAGCATCTTTGCAAGTCAAGAA 1459
Db 1357 AAGAGCTTACAGAGCATATTTAAAAATCTTCAGACTCAACAAGCATCTTTGCAAGTCAAGAA 1416
Qy 1460 ACACACATTTTGAATAGAAATCAGAACTTCAACAGAAAATTTAAAGTAATGACTGAAT 1519
Db 1417 ACACACATTTTGAATAGAAATCAGAACTTCAACAGAAAATTTAAAGTAATGACTGAAT 1476
Qy 1520 TATATCAAGAAATGAATGAABACTCCACAGGAAATTTAACAGTAGAGGAAAATTTATCGGT 1579
Db 1477 TATATCAAGAAATGAATGAABACTCCACAGGAAATTTAACAGTAGAGGAAAATTTATCGGT 1536
Qy 1580 TAGAGAAAGAGAGAACTTTCTAAAGTAGATGAAGAAATCAGCCATGCGACCTGAAGAGC 1639
Db 1537 TAGAGAAAGAGAGAACTTTCTAAAGTAGATGAAGAAATCAGCCATGCGACCTGAAGAGC 1596
Qy 1640 TGGAGCCTATAGAAAGCGGACCAAGATCTTGAAGAAATTTGAAGAGAACTTATTCAT 1699
Db 1597 TGGAGCCTATAGAAAGCGGACCAAGATCTTGAAGAAATTTGAAGAGAACTTATTCAT 1656
Qy 1700 CTTATCAAGGGCAGATTTATTTCCATGAGAAAAGCACAATGATAAATGGTTGGCAGCTC 1759
Db 1657 CTTATCAAGGGCAGATTTATTTCCATGAGAAAAGCACAATGATAAATGGTTGGCAGCTC 1716
Qy 1760 GGAATGCTGAAAGAACTCAATGATTTAAGGAAGAAATGCTCAACAGACACAAAAT 1819
Db 1717 GGAATGCTGAAAGAACTCAATGATTTAAGGAAGAAATGCTCAACAGACACAAAAT 1776
Qy 1820 TAACCTGAAACAGAGCTTTAAATTTGAACTTTTGAAGAAAGATCCTTATGCACTCGATGTTT 1879
Db 1777 TAACCTGAAACAGAGCTTTAAATTTGAACTTTTGAAGAAAGATCCTTATGCACTCGATGTTT 1836
Qy 1880 CAAATACAGATTTGCGAGAGCATTTCCCATATGGTCCCTCACCATTGGGTTGGCCTT 1939
Db 1837 CAAATACAGATTTGCGAGAGCATTTCCCATATGGTCCCTCACCATTGGGTTGGCCTT 1896
Qy 1940 CATCTGAAAACAAGAGCTTTTCTCTCTCTCCAACTTTGTTGGAGGTTCCACTCAGACTCT 1999
Db 1897 CATCTGAAAACAAGAGCTTTTCTCTCTCTCCAACTTTGTTGGAGGTTCCACTCAGACTCT 1956
Qy 2000 CACCTTTGCTTTCAGGGGAGGAGAGAGCTCAAGCTGTGATAGTTAAACCGATCCTCATAGG 2059
Db 1957 CACCTTTGCTTTCAGGGGAGGAGAGAGCTCAAGCTGTGATAGTTAAACCGATCCTCATAGG 2016
Qy 2060 ATCAGATTTCCAAATGAAGAGGAGAAATCAAGCTGTGATAGTTAAACCGATCCTCATAGG 2119
Db 2017 ATCAGATTTCCAAATGAAGAGGAGAAATCAAGCTGTGATAGTTAAACCGATCCTCATAGG 2076
Qy 2120 CTCCTCTGACACATGGGTCTCTGTCCTCATGGGACAGGACCGTAGATGATGTTTTC 2179
Db 2077 CTCCTCTGACACATGGGTCTCTGTCCTCATGGGACAGGACCGTAGATGATGTTTTC 2136
Qy 2180 CTCGGCAGAGCAATCATATCTGATTCAGCCCTTCTCCACAAAGCAAGACAGATTTT 2239
Db 2137 CTCGGCAGAGCAATCATATCTGATTCAGCCCTTCTCCACAAAGCAAGACAGATTTT 2196
Qy 2240 GTTCTTAATTTGTTAGACTGTTCTGACACAGCAAGAACTCAGAAATTTTAAATATGCTTCTT 2299
Db 2197 GTTCTTAATTTGTTAGACTGTTCTGACACAGCAAGAACTCAGAAATTTTAAATATGCTTCTT 2256
Qy 2300 TGGATAAATGATGGGTCAATGCTTCCAGAAATGGAATCCAGTAGAAATGATACCAAG 2359
Db 2257 TGGATAAATGATGGGTCAATGCTTCCAGAAATGGAATCCAGTAGAAATGATACCAAG 2316
Qy 2360 ATGATCTTTGGTAAATTTAAATGCTGCTGATTCATCTCTCCCTGCTGAAAATGAAGCCACTG 2419
Db 2317 ATGATCTTTGGTAAATTTAAATGCTGCTGATTCATCTCTCCCTGCTGAAAATGAAGCCACTG 2376

2420	GCCTGGCTTTGTTCTCTCCACCTCTTGCTCCAAATCAGAGGTCCATTTGTTTCCAGTGGATG	2479
2377	GCCTGGCTTTGTTCTCTCCACCTCTTGCTCCAGTCCAGAGGTCCATTTGTTTCCAGTGGATG	2436
2480	CAAGAGGCCCATTTCTTGAGAGGAGGACCTCTTTCCGCCCACTCTCTCCAGG-AGCCATG	2538
2437	CAAGAGGCCCATTTCTTGAGAGGAGGACCTCTTTCCGCCCACTCTCTCCAGGAGCCATG	2496
2539	TTTGGAGCTTCTCGAGATTATTTTCCACCAAGG---ATTTCAGGTCCACCACTCC-TCG	2594
2497	TTTGGAGCTTCTCGAGATTATTTTCCACCAAGGAGATTTCACAGGGGATTTCACCACTTGC	2556
2595	TCATTG---CAATGAGAAATGTCTATCCACCGAGGGGTTTTCCTCT---TACCT	2645
2557	TCATTGGCAATGAGGAAATGCTTATCCACCGAGGGGTTTTCCTCTTAACTT	2616
2646	TCCCCCAAGACCT---GGATTTTCCCCCACTCCACCCACATCTC--GAAAGTAAAGTGAG	2700
2617	TCCCCCAAGACNGGGGTTTTTCCCCCACTCCACCCACATTTTGAAGGTAGAGAGTGAA	2676
2701	TTCCCTCTCAGG-TTTGATTCCACCTTCAAAATGAGCCTGTCTACTGAACATCCAGAACCA	2759
2677	TTTCCCTTAGGTTTGTGATTCCAACTTCAAAATGAGCCTGTCTACTGAACATCCAGAACCA	2736
2760	GCAAG-AAACCTGACAATATTTT-GCTCTCTTCAAAGTAATTTTGACCTGATCTCATTT	2817
2737	GCAAGAAACCTGACAATATTTTGGCTCTCTTCAAAGTAATTTTGACCTGATCTCATTT	2796
2818	TCAGTTAACTAACTGCTGTACTTAAAGTGATTACACTTTT-GCTCAAATGGAAGCTTAA	2876
2797	TCAGTTAACTAACTGCTGTACTTAAAGTGATTACACTTTTGGCTCAAATGGAAGCTTAA	2856
2877	TGCAATTAATCTCTCAGGATAGTATTTTGTAAATAAGATGATTTTAAATATGAATCTTA	2936
2857	TGCAATTAATCTCTCAGGATAGTATTTTGTAAATAAGATGATTTTAAATATGAATCTTA	2916
2937	TGAGTAAATTTTCAATTTTATTTTGAACGGTAACTATTTTCAATTTGATTAATCCAC	2996
2917	TGAGTAAATTTTCAATTTTATTTTGAACGGTAACTATTTTCAATTTGATTAATCCAC	2976
2997	TATATATAACAACTAGTGGAGTTTATATATCTATCTATCTTTCAGGTGGGAGGCTTAA	3056
2977	TATATATAACAACTAGTGGAGTTTATATATCTATCTATCTTTCAGGTGGGAGGCTTAA	3036
3057	ATTCGAGTCTGTCTTTATGCAAGAACTGTATTTTCTGTGGTGTGGACAAATGTG	3116
3037	ATTCGAGTCTGTCTTTATGCAAGAACTGTATTTTCTGTGGTGTGGACAAATGTG	3096
3117	AAAGCTACTTTATCTTAAATAAATATAGTTGATTTAAAAA	3158
3097	AAAGTAACTTTATCTTAAATAAATATAGTTGATTTAAAGA	3138
RESULT 4		
ABQ77424		
ID	ABQ77424 standard; cDNA; 2928 BP.	
XX	AC ABQ77424;	
XX	10-MAY-2003 (first entry)	
XX	Human CGDD cDNA 6715627CB1 SEQ ID 38.	
XX	Human; cell growth, differentiation and death; CGDD; cardiact; cancer;	
XX	cytostatic; neuroprotective; gene therapy; diagnosis; cardiovascular;	
XX	neurological disorder; gene; ss.	
OS	Homo sapiens.	
XX	WO2003014322-A2.	
XX	20-FEB-2003.	

08-AUG-2002; 2002WO-US025465.
08-AUG-2001; 2001US-0311017P.
17-AUG-2001; 2001US-0313070P.
17-AUG-2001; 2001US-0313071P.
24-AUG-2001; 2001US-0314678P.
31-AUG-2001; 2001US-0316692P.
07-SEP-2001; 2001US-0317913P.
14-SEP-2001; 2001US-0322182P.
07-DEC-2001; 2001US-0340747P.
20-DEC-2001; 2001US-0342761P.
29-MAR-2002; 2002US-0369129P.
(INCY-) INCYTE GENOMICS INC.
Azimzai Y, Barroso I, Baughn MR, Becha SD, Borowsky ML; Duggan BM, Elliott VS, Emmerling BM, Forsythe IJ, Gietzen KJ; Gorvad AE, Graul RC, Griffin JA, Gururajan R, Hafalia AJA, Ison CH; Kable AE, Khan FA, Lee S, Lee SY, Li JX, Reddy R, Richardson TW; Sprague WW, Swarnakar A, Tang YT, Warren BA, Xu Y, Yao MG, Yue H; Yue H;
WPI; 2003-268197/26.
P-PSDB; ABG74685.
New polypeptide associated with cell growth, differentiation and death, useful for preparing a composition for diagnosing or treating a disease e.g., cardiovascular or neurological disorders or cancer.
Claim 93; Page 251-252; 267pp; English.
This invention describes a novel polypeptide associated with cell growth, differentiation and death (CGDP) which has cardiant, cytostatic and neuroprotective activity. The polypeptides of the invention are useful for preparing a composition for diagnosing, or treating a disease or condition associated with decreased expression or overexpression of functional CGDP e.g., cardiovascular or neurological disorders or cancer. The polynucleotides of the invention can be used for gene therapy. CC ABQ77414-ABQ77441 encode the human CGDP polypeptides represented in CC ABG74675-ABG74701
XX Sequence 2928 BP; 909 A; 597 C; 660 G; 762 T; 0 U; 0 Other;
Query Match 79.5%; Score 2515.2; DB 7; Length 2928;
Best Local Similarity 93.6%; Pred.No. 0;
Matches 2708; Conservative 0; Mismatches 158; Indels 28; Gaps 7
QY 90 CTAAGGGGGGCCGGCTCAGACAGCGCTGCCTCAGATGAAGTGTAAACAAGAGGGCC 149
Db 35 CTGTGCGGGTCGGGCTCGGACTTGCGCTGCCCTCGGGATGAAGTATAACAAGAGGGTC 94
QY 150 AGGGAGGTGTGGGGGACAAATGGSCGTGTAGGCTGTGGTCCCGCGTTCGCCAG 209
Db 95 GGATGTGTACGTAGCGCTGTGTAGCCCTGT-----CGGTGCCCCCTGTCCCACATCG 148
QY 210 CTCCTCCCGCAGCCCGCTCCACAGTGTTCGCTTCGGTGTGTGTCACTGCGCATTCGG 269
Db 149 TCCCTCCAGCAGCAGCGGCTCCGCACTGCTCCACTCCGCT----TCCCGGGTGGGATTCCG 204
QY 270 GTTCAGACCACAGCTGCGTGTCTCCACCGTTGTGT-----GGCCAGTGT 319
Db 205 GTTCGGACCGAAGCTGTGTGTCTCCCGGTATTGTGGCCCCCAGCGCCGGGTT 264
QY 320 ACTCGGTGACCGCACAGCAGCTCGACGATATGGAGAGCTGTGTGTACCTCCCTCAGC 379
Db 265 ACTGTGGCGACCAAGACAGCAGCTTTCGCGTATGGAGGAGCCCGGGGCTTACCTCTAAC 324
QY 380 CCTACTCGGGCTGTCTCTGGAGAGCTTACGACAGTGTGGCAGACTACTCTGAGAGTA 439
Db 325 CGTATTGGGCTGTCTCTGGAGAGCTTACGACAGTGTGGCAGACTCTGCTGAAGTA 384
QY 440 TGAGCCAGATAGAAATCCTTAGGTTTTTCATCGGAATCTGTTGTTATGTGAGCTGTTA 499

GenCore version 5.1.6
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2004, 02:11:04 ; Search time 212 Seconds
(without alignments)
8282.380 Million cell updates/sec

Title: US-09-744-314-14

Perfect score: 3164

Sequence: 1 aggttaacatcatgaagag.....agttgatttaaaaaaaaaa 3164

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/2/ina/5A-COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B-COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A-COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B-COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PTUS-COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	3164	100.0	3164	US-09-744-314-14
2	2416.2	76.4	2770	US-09-744-314-14
3	78	2.5	116592	US-09-744-314-14
4	70.4	2.2	7218	US-09-744-314-14
5	59.8	1.9	3095	US-09-744-314-14
6	56.2	1.8	11131	US-09-744-314-14
7	54.4	1.7	636	US-09-744-314-14
8	52.6	1.7	1664976	US-09-744-314-14
9	52.6	1.7	1664976	US-09-744-314-14
10	51.6	1.6	837	US-09-744-314-14
11	51.2	1.6	19124	US-09-744-314-14
12	51	1.6	1664976	US-09-744-314-14
13	50.8	1.6	2637	US-09-744-314-14
14	50	1.6	3222	US-09-744-314-14
15	49.2	1.6	6306	US-09-744-314-14
16	49	1.5	832	US-09-744-314-14
17	49	1.5	6326	US-09-744-314-14
18	49	1.5	168575	US-09-744-314-14
19	48.8	1.5	30549	US-09-744-314-14
20	48	1.5	580073	US-09-744-314-14
21	47.8	1.5	8607	US-09-744-314-14
22	47.4	1.5	665	US-09-744-314-14
23	47.4	1.5	6156	US-09-744-314-14
24	47	1.5	10640	US-09-744-314-14
25	46.8	1.5	2169	US-09-744-314-14
26	46.6	1.5	6866	US-09-744-314-14
27	46.4	1.5	9347	US-09-744-314-14

28	46.4	1.5	640681	4	US-09-790-988-1	Sequence 1, Appli
29	46.2	1.5	6306	4	US-10-204-708-50	Sequence 50, Appl
30	46.2	1.5	7664	4	US-10-204-708-84	Sequence 84, Appl
31	45.8	1.4	3446	4	US-09-392-714-13	Sequence 13, Appl
32	45.6	1.4	3636	4	US-09-134-001C-1983	Sequence 86, Appl
33	45.2	1.4	11050	4	US-10-204-708-86	Sequence 86, Appl
34	45	1.4	2017	4	US-09-404-879A-72	Sequence 72, Appl
35	45	1.4	2017	4	US-09-338-933-72	Sequence 72, Appl
36	45	1.4	2017	4	US-09-215-681-72	Sequence 72, Appl
37	45	1.4	2017	4	US-09-215-003A-72	Sequence 72, Appl
38	44.4	1.4	633	4	US-09-134-001C-578	Sequence 578, App
39	44.4	1.4	1707	4	US-09-134-001C-931	Sequence 931, App
40	44.4	1.4	4766	5	PCT-US93-07261-10	Sequence 10, Appl
41	44.4	1.4	5433	3	US-08-929-329-11	Sequence 11, Appl
42	44.2	1.4	6317	4	US-10-204-708-11	Sequence 11, Appl
43	44.2	1.4	19124	2	US-08-487-826B-13	Sequence 13, Appl
44	44.2	1.4	19233	4	US-10-204-708-46	Sequence 46, Appl
45	44.2	1.4	20674	4	US-09-641-638-651	Sequence 651, App

ALIGNMENTS

RESULT 1

US-09-976-594-320

Sequence 320, Application US/09976594

Patent No. 6673549

GENERAL INFORMATION:

APPLICANT: Furness, Michael

APPLICANT: Buchbinder, Jenny

TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS

FILE REFERENCE: PA-0041 US

CURRENT APPLICATION NUMBER: US/09/976,594

CURRENT FILING DATE: 2001-10-12

PRIOR APPLICATION NUMBER: 60/240,409

PRIOR FILING DATE: 2000-10-12

NUMBER OF SEQ ID NOS: 1143

SOFTWARE: PERL Program

SEQ ID NO 320

LENGTH: 3164

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No. 6673549 3768043CB1

US-09-976-594-320

Query Match 100.0%; Score 3164; DB 4; Length 3164;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	AGGTTTATCATGAAGAAGACAGCAATTTTAAAGTGTTTACCAAAATAAAGCTT	60
DB	1	AGGTTTATCATGAAGAAGACAGCAATTTTAAAGTGTTTACCAAAATAAAGCTT	60
QY	61	CAAAATATGTGATGTGAAACTGCGAGAACTAAGCGCGCGGCTCAGACCGGCTGC	120
DB	61	CAAAATATGTGATGTGAAACTGCGAGAACTAAGCGCGCGGCTCAGACCGGCTGC	120
QY	121	CTCAGATGTAAAGTGTAACAGAGGCGCGGAGTGTTGGGGCAACATGGGCTG	180
DB	121	CTCAGATGTAAAGTGTAACAGAGGCGCGGAGTGTTGGGGCAACATGGGCTG	180
QY	181	TGAGGCGCTGTGGGTGCGCGCGTTCGCCAGCTCCCGCCAGCCCGCTCCACAGTGTGCG	240
DB	181	TGAGGCGCTGTGGGTGCGCGCGTTCGCCAGCTCCCGCCAGCCCGCTCCACAGTGTGCG	240
QY	241	CTCCGGTGTGTTGTCAGTGGCGATTCGGTTCAGAGCCAGGCTGGTGTCTCCACC	300
DB	241	CTCCGGTGTGTTGTCAGTGGCGATTCGGTTCAGAGCCAGGCTGGTGTCTCCACC	300
QY	301	GCTTGTGTGGCCAGTGTTACTTGCCTGACCGCCAGCAGCCTCGAGCTATGGAGAG	360

320
Claims
7 DWIPI

Db 301 GCTTGTGTGGCCAGTGTACTGCGGTGACCGCCAGAGCAGCTCGACGCTATGGAGGAG 360
Qy 361 CTTGTGTGTACCCCTCAGCCCTACCTGGGCTGGTCTGGAGGAGCTACGAGAGTGTG 420
Db 361 CTTGTGTGTACCCCTCAGCCCTACCTGGGCTGGTCTGGAGGAGCTACGAGAGTGTG 420
Qy 421 GCAGCACTACCTGAGAGTATGAGACAGATGAGAACTCTTATGGTTTCCATCGAACTG 480
Db 421 GCAGCACTACCTGAGAGTATGAGACAGATGAGAACTCTTATGGTTTCCATCGAACTG 480
Qy 481 GTGTGTATGTCAGCTGTATTCGATTTTGTGTGTCTCTTTTGTGTGTGTGTGTGTGT 540
Db 481 GTGTGTATGTCAGCTGTATTCGATTTTGTGTGTCTCTTTTGTGTGTGTGTGTGTGT 540
Qy 541 AGATCGGTGAGAGTGGCTTTACCTGGGAGAGAGCAAACTTGTGTGCAACGCTTCT 600
Db 541 AGATCGGTGAGAGTGGCTTTACCTGGGAGAGAGCAAACTTGTGTGCAACGCTTCT 600
Qy 601 GCACCTAATTTGAAGAAATGTAATCTTGAAGAAATTTAGCCCTTATTCAAAGAGAT 660
Db 601 GCACCTAATTTGAAGAAATGTAATCTTGAAGAAATTTAGCCCTTATTCAAAGAGAT 660
Qy 661 GAAGCTATGAAGTATGAGTCACTTTTAGAGGATGCCAGCTTTGAGAGGAGGAGCAAA 720
Db 661 GAAGCTATGAAGTATGAGTCACTTTTAGAGGATGCCAGCTTTGAGAGGAGGAGCAAA 720
Qy 721 GAAGCAGAGTGTGGAGCACTGTGAAAGCTGAAAGCTCAATTTCTGAACCTTGAG 780
Db 721 GAAGCAGAGTGTGGAGCACTGTGAAAGCTGAAAGCTCAATTTCTGAACCTTGAG 780
Qy 781 GATGAATCTCTGTCTAGAAAAAGACTTTAAACCAAGAGAAATCTTCAACAA 840
Db 781 GATGAATCTCTGTCTAGAAAAAGACTTTAAACCAAGAGAAATCTTCAACAA 840
Qy 841 GATGAATGATGGGGATATTTCAAAAGTATACAGTCTTAGAAGATGAGTCAAAATCC 900
Db 841 GATGAATGATGGGGATATTTCAAAAGTATACAGTCTTAGAAGATGAGTCAAAATCC 900
Qy 901 CTCAAATCAAAATAGCTGAAGCAAAATCACTCTGCAAGACATTTAAATGAGTGAAGAA 960
Db 901 CTCAAATCAAAATAGCTGAAGCAAAATCACTCTGCAAGACATTTAAATGAGTGAAGAA 960
Qy 961 CGACGGCTATAGCAATAAAGATGCTTTGAATGAAATTTCTCACTCAGACAGCCAT 1020
Db 961 CGACGGCTATAGCAATAAAGATGCTTTGAATGAAATTTCTCACTCAGACAGCCAT 1020
Qy 1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGGAAGAGGAGAGTGAAGTGAATTAATAA 1080
Db 1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGGAAGAGGAGAGTGAAGTGAATTAATAA 1080
Qy 1081 CAGAAATAAATTTGAAGACTCCAAAGTACACGAGAACAAAGTTCTGAATGATAAGAA 1140
Db 1081 CAGAAATAAATTTGAAGACTCCAAAGTACACGAGAACAAAGTTCTGAATGATAAGAA 1140
Qy 1141 AATCATCATCAAGCCCTGACTGACACTGCACTGCAATGCAATGCAATGCAATGCAAT 1200
Db 1141 AATCATCATCAAGCCCTGACTGACACTGCACTGCAATGCAATGCAATGCAATGCAAT 1200
Qy 1201 GAAGAGACACAAACCGATGATGATACCTGGAATTTAGAGTGAACAGTGAATCGGAAAT 1260
Db 1201 GAAGAGACACAAACCGATGATGATACCTGGAATTTAGAGTGAACAGTGAATCGGAAAT 1260
Qy 1261 GGTGCTTACTAGATTAATCTTCAAAAGAGCTTTGAAGAACTGATTCATGCTGCTAAG 1320
Db 1261 GGTGCTTACTAGATTAATCTTCAAAAGAGCTTTGAAGAACTGATTCATGCTGCTAAG 1320
Qy 1321 TTAATGCTTTCTTTAAACCTTTAGAGGAGAGAAACCAATTTATATTCACTGTGCT 1380
Db 1321 TTAATGCTTTCTTTAAACCTTTAGAGGAGAGAAACCAATTTATATTCACTGTGCT 1380
Qy 1381 GAAGTTGATAAACAAGAGAGAGCTTACAGAGCATATTAATAATCTTCAGACTCAACAA 1440
Db 1381 GAAGTTGATAAACAAGAGAGAGCTTACAGAGCATATTAATAATCTTCAGACTCAACAA 1440

Qy 1441 GCATCTTTTCAGTCAAGAAACACACATTTTGAATTCAGAACTCAGAGCTTCAACAGAAA 1500
Db 1441 GCATCTTTTCAGTCAAGAAACACACATTTTGAATTCAGAACTCAGAGCTTCAACAGAAA 1500
Qy 1501 CTTAAAGTAAATGACTGAATTTATCAAGAAATGAAATGAAATGAAATGAAATGAAATGAA 1560
Db 1501 CTTAAAGTAAATGACTGAATTTATCAAGAAATGAAATGAAATGAAATGAAATGAAATGAA 1560
Qy 1561 GTAGAGGAAATTTATCGGTTTAGAGAAAGAGAGAACTTTTAAAGTAGATGAAAGATC 1620
Db 1561 GTAGAGGAAATTTATCGGTTTAGAGAAAGAGAGAACTTTTAAAGTAGATGAAAGATC 1620
Qy 1621 AGCCATCCCACTGAAGAGCTGGAGACTATAGAAAGCGAGCAAAAGATCTTGAAGAGAA 1680
Db 1621 AGCCATCCCACTGAAGAGCTGGAGACTATAGAAAGCGAGCAAAAGATCTTGAAGAGAA 1680
Qy 1681 TTGGAGAGAACTATTCTTCTTATCAAGGCGAGATTTTCCATGAGAGAAAGACAT 1740
Db 1681 TTGGAGAGAACTATTCTTCTTATCAAGGCGAGATTTTCCATGAGAGAAAGACAT 1740
Qy 1741 GATAATTTGGTTGCGAGCTCGGAATGCTGAAAGAAACCTCAATGATTTAAGGAAAGAAAT 1800
Db 1741 GATAATTTGGTTGCGAGCTCGGAATGCTGAAAGAAACCTCAATGATTTAAGGAAAGAAAT 1800
Qy 1801 GCTCAAAACAGCAAAATTTAACTGAAACAGAGCTTAAATTTGAACTTTTGAAGAAAGAT 1860
Db 1801 GCTCAAAACAGCAAAATTTAACTGAAACAGAGCTTAAATTTGAACTTTTGAAGAAAGAT 1860
Qy 1861 CTTTATGCACTCGATTTTCCAAATACAGATTTTGGCAGAGAGCATTTCCCATATGTCCTC 1920
Db 1861 CTTTATGCACTCGATTTTCCAAATACAGATTTTGGCAGAGAGCATTTCCCATATGTCCTC 1920
Qy 1921 TCACCAATTTGGTTGGCTTCACTGAAACAGAGCTTTTCTCTCTCTCAACTTTGTTG 1980
Db 1921 TCACCAATTTGGTTGGCTTCACTGAAACAGAGCTTTTCTCTCTCTCAACTTTGTTG 1980
Qy 1981 GAGGCTCCACTCAGACTCTCACCTTTTCTTCCAGGGGAGGAGAGAGGCTCACAGGC 2040
Db 1981 GAGGCTCCACTCAGACTCTCACCTTTTCTTCCAGGGGAGGAGAGAGGCTCACAGGC 2040
Qy 2041 CAGAGGAACTCTGACACATCAGATTTACGATGAAAGAGGAGAACTCAAGCTGTGATAG 2100
Db 2041 CAGAGGAACTCTGACACATCAGATTTACGATGAAAGAGGAGAACTCAAGCTGTGATAG 2100
Qy 2101 TTAACGATCTCTATAGGCTCTCTGACACTGAGCTTCTGTCACTCCATCCATGGGACCA 2160
Db 2101 TTAACGATCTCTATAGGCTCTCTGACACTGAGCTTCTGTCACTCCATCCATGGGACCA 2160
Qy 2161 GACCGTAGGATGATTTCTCTCGCAGGACAAATCATATCTCTGATTCAGCCCTTCTCTCA 2220
Db 2161 GACCGTAGGATGATTTCTCTCGCAGGACAAATCATATCTCTGATTCAGCCCTTCTCTCA 2220
Qy 2221 CAAAGCAGACAGATTTTGTCTTAATTTCTGGTAGACTGTCTGGACCCAGCAACTCAGA 2280
Db 2221 CAAAGCAGACAGATTTTGTCTTAATTTCTGGTAGACTGTCTGGACCCAGCAACTCAGA 2280
Qy 2281 AGTTTAAATGCTCTTTTGGATAAAATGAGTGGGTCAATGCCCTTCAAGAAATGGAATCC 2340
Db 2281 AGTTTAAATGCTCTTTTGGATAAAATGAGTGGGTCAATGCCCTTCAAGAAATGGAATCC 2340
Qy 2341 AGTAGAAATGATACCAAGATGATCTTGGTAAATTTAAATGTCCTGATTCATCTCTCCCT 2400
Db 2341 AGTAGAAATGATACCAAGATGATCTTGGTAAATTTAAATGTCCTGATTCATCTCTCCCT 2400
Qy 2401 GCTGAAATGAAGCCACTGGCCCTTGTCTTCTCCACCTTCTCTCAATTCAGAGGT 2460
Db 2401 GCTGAAATGAAGCCACTGGCCCTTGTCTTCTCCACCTTCTCTCAATTCAGAGGT 2460
Qy 2461 CCATTTGTTCCAGTGAAGAGAGCCCTTCTGAGAGAGAGCTCTCTTCCCTCCCA 2520
Db 2461 CCATTTGTTCCAGTGAAGAGAGCCCTTCTGAGAGAGAGCTCTCTTCCCTCCCA 2520

QY 1379 CTGAAGTTGATAAACAAGAGAGAGCTTACAGACATATTAAATACTTCAGACTCAAC 1438
DB 1142 ATGAATAGTAGAAGTAGAAGAAATACCTTCAGAGAGATGATAAATGAAAAAGTTGAC 1201
QY 1439 AAGCATCTTTGAGTCAGAGAAACACACATTTTGAATAATGAGAAATCAGAGCTTCAACAGA 1498
DB 1202 ATGAATAATAGTAGAAGTTGAAGAAATCTTACCAGAGATATAAATAATGAAAAAGGTCACATG 1261
QY 1499 AACTTAAAGTAATGACTGAAATATATCAAGAAATGAAATGAAATCCACAGGAAATTA 1558
DB 1262 AATAGTAGAGGTTGAAGAAATCTTACCAGAGATGATAAATGAAAAAGTTGACATG 1321
QY 1559 CAGTAGAGAAATATATCGGTTAGAGAAAGAGAGAAACTTTCTAAATAGATGAGAAAGA 1618
DB 1322 AATAGTAGAGGTTGAAGAAATCTTACCAGAGATATAAATAATGAAAAAGGTCACATGAA 1381
QY 1619 TCAGCCATCCCACTGAAGAGCTGGAGACCTATAGAAAGCGAGCCAAAGATCTTGAAGAG 1678
DB 1382 TAGTAGAGGTTGAAGAAATCTTACCAGAGATATAAATAATGAAAAAGTTGACATGAAATAG 1441
QY 1679 AATGGAGAGAACTATTCATCTTATCAAGGCGAGATTTTCCCATGAGAAAAAGCAC 1738
DB 1442 TAGAAGTTGAAGAAATCTTACCAGAGATATAAATAATGAAAAAGGTCACATGAAATAGTAG 1501
QY 1739 ATGATAATTTGGTGGCGAGCTCGAATCTGAGAGAACTCAATGATTTAGGAAAGAA 1798
DB 1502 AGTTGAAGAAATCTTACCAGAGATATAAATAATGAAAAAGTTCAACATGAAATAGTAGAG 1561
QY 1799 ATGCTCACACAGACAAAAATTAACATGAAAGAGCTTAAATTTGAACCTTTTGA 1853
DB 1562 TTGAAGAAATCTTACCAGAGATATAAATAATGAAAAAGGTCACATGAAATAGTAGA 1616

RESULT 6

US-10-204-708-28
; Sequence 28, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2003-05-06
; PRIOR FILING DATE: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; LENGTH: 11131
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-28

Query Match 1.8%; Score 56.2; DB 4; Length 11131;
Best Local Similarity 46.2%; Pred. No. 7.8e-05;
Matches 187; Conservative 0; Mismatches 218; Indels 0; Gaps 0;
QY 2744 AACATCCAGAACACACAGCAGAAACCTGACAAATATTTTCTCTCTTCAAAAGTAATTTT 2803
DB 10465 AAAAATTAATGAAGAAAGAAATTAAGAGTTAGTAGTATTATTTTATTTAGAGTTGAATATT 10524

QY 2804 GACTGATCTCATTTTCAGTTTAAAGTAACCTGCTTACTTAAGTATTACACTTTTGCTCA 2863
DB 10525 TAAGGAABATAGGTATGGTTTATTGATTTATTTATTAGATAATATAAATTTTGAAGT 10584
QY 2864 AATTGAAGCTTAATGAATTAATATTCACAGATAGTATTTTGTAAATAAGATGATTTA 2923
DB 10585 TGGGTATGTTTATGAGTTTATAAGTATGTTATATATTGATTTTATTTTAAATATATTG 10644
QY 2924 AATATCAATCTTATGAGTAAATTTATTTCAATTTTATTTAGACGGTATATATTTCAAT 2983
DB 10645 AATTTTAAAAAATTAATTTGTTTTTATTTATGATGAATTTTGAAGTATATTAAT 10704
QY 2984 TTGATTAATCCACTATATATAAACAATAGTGGAGTTTATATATATGTAATCTTCAGGT 3043
DB 10705 AGTTATTTATTTATGAGATTTATATGTTTAGATTTTATTTTGTAGTATTTTGGCT 10764
QY 3044 GGGGAGGCTTTAAATTTCTGAAGTCTGCTTTATGCCAAGAACTGTTATTTACTGTGGTT 3103
DB 10765 GTATTATAATTTTATATATAATTTTGTGAGATAAGTTTGTATTGTTTATTTTGTG 10824
QY 3104 GTGACAAATGTGAAGTAACTTTATGCTTAAATAAATATATAGTT 3148
DB 10825 ATTGAAGAAATTTGAGAAGTAGAGTGTAAAGAAATTTGTTTATAGT 10869

RESULT 7

US-08-998-416-1137
; Sequence 1137, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgien
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPHII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC-976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 1137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.
OM nucleic - nucleic search, using sw model
Run on: April 21, 2004, 02:53:20 ; Search time 1286 Seconds
(without alignments)
11092.765 Million cell updates/sec

Title: US-09-744-314-14
Perfect score: 3164
Sequence: 1 aggttaaccatgaagaag.....agttgatttaaaaaaaaaa 3164

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2907579 seqs, 2254313464 residues

Total number of hits satisfying chosen parameters: 5815158

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA.*

- 1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq2.*
- 14: /cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq.*
- 16: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq.*
- 17: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq.*
- 18: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq.*
- 19: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2414.6	76.3	2814	9	US-09-880-107-3419
2	2414.6	76.3	2814	9	US-09-880-107-3419
3	2414.6	76.3	2814	13	US-10-081-056-335
4	2414.6	76.3	2814	13	US-10-305-654-335
5	2414.6	76.3	2814	15	US-10-223-085-335
6	2414.6	76.3	2814	15	US-10-223-084-335
7	2414.6	76.3	2814	15	US-10-223-088-335
8	2414.6	76.3	2814	15	US-10-223-089-335
9	2414.6	76.3	2814	15	US-10-223-087-335
10	2414.6	76.3	2814	15	US-10-223-083-335
11	2414.6	76.3	2814	15	US-10-223-089-335
12	2414.6	76.3	2814	15	US-10-223-081-335
13	2414.6	76.3	2814	15	US-10-223-082-335
14	2397.2	75.8	2602	16	US-10-161-927-27

15	2303.4	72.8	2614	16	US-10-161-927-29	Sequence 29, Appl	
16	2159.8	68.3	2351	9	US-09-864-761-21534	Sequence 21534, A	
C	17	2148.6	67.9	2334	9	US-09-864-761-30648	Sequence 30648, A
18	2111.4	66.7	2588	16	US-10-161-493-53	Sequence 53, Appl	
19	2026.2	64.0	2483	16	US-10-161-493-55	Sequence 55, Appl	
20	1954.6	61.8	4292	9	US-09-729-674-41	Sequence 41, Appl	
21	1838.8	58.1	1986	9	US-09-864-761-4795	Sequence 4795, Ap	
C	22	1797.8	56.8	1961	9	US-09-864-761-14090	Sequence 14090, A
23	912.8	28.8	925	15	US-10-029-386-20575	Sequence 20675, A	
C	24	871.8	27.6	883	15	US-10-029-386-20733	Sequence 20733, A
25	777.2	24.6	1044	9	US-09-822-830A-596	Sequence 596, App	
26	565.8	17.9	569	15	US-10-029-386-6964	Sequence 6964, Ap	
C	27	562.6	17.8	569	15	US-10-029-386-7023	Sequence 7023, Ap
C	28	559.8	17.7	643	9	US-09-920-300A-163	Sequence 163, App
C	29	559.8	17.7	643	14	US-10-033-528-163	Sequence 163, App
C	30	559.8	17.7	643	15	US-10-099-926-163	Sequence 163, App
C	31	479.8	15.2	559	9	US-09-864-761-12319	Sequence 12319, A
32	466	14.7	597	13	US-10-027-632-265229	Sequence 265229, A	
33	466	14.7	597	16	US-10-027-632-265229	Sequence 265229, A	
34	464.4	14.7	597	13	US-10-027-632-265228	Sequence 265228, A	
35	464.4	14.7	597	16	US-10-027-632-265228	Sequence 265228, A	
C	36	462	14.6	593	13	US-10-027-632-264782	Sequence 264782, A
C	37	462	14.6	593	16	US-10-027-632-264782	Sequence 264782, A
C	38	460.4	14.6	593	13	US-10-027-632-264783	Sequence 264783, A
C	39	460.4	14.6	593	16	US-10-027-632-264783	Sequence 264783, A
40	458.8	14.5	543	15	US-10-060-036-1501	Sequence 1501, Ap	
C	41	438.4	13.9	440	9	US-09-998-598-1014	Sequence 1014, Ap
C	42	436.4	13.8	480	13	US-10-027-632-268158	Sequence 268158, A
C	43	436.4	13.8	480	16	US-10-027-632-268158	Sequence 268158, A
C	44	434.8	13.7	480	13	US-10-027-632-268157	Sequence 268157, A
C	45	434.8	13.7	480	13	US-10-027-632-268159	Sequence 268159, A

ALIGNMENTS

RESULT 1

US-09-880-107-3419
; Sequence 3419, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3419
; LENGTH: 2814
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U73682
US-09-880-107-3419

Query Match	76.3%	Score 2414.6	DB 9	Length 2814
Best Local Similarity	92.3%	Pred. No. 0		
Matches 2630	Conservative	0	Mismatches 84	Indels 135
Gaps				3
QY	310	GCACAGTGTACTGCGGTGACCGCAGACGCTCGACGCTATGAGAGACCTGTGCT	369	
Db	5	GCACGGGTGTACTGTGGCACCACAGAGACGCTTTGGCGCTATGAGAGACCTGTGCT	64	
QY	370	ACCCTCAGCCCTACTGCGGTGCTGCTGAGGAGCTACGAGAGTGTGCGACACTA	429	

Db 65 ACCCTCAACCGTATTGGGGTCTCTCGAGGAGCTACGACGGTTTGGCCAGCACTG 124
QY 430 CTGAGAGTATGAGACCACTGAGAACTCTTATGTTTTCATCGGAATCGTGTGTATGT 489
Db 125 CTGAGGTATGAGACCACTGATCTTATCTTTATGTTTTCATCGGAATCGTGTATGT 184
QY 490 GAGCTGTTATGGAATTTTGTGTTTCTCTTTTGTGAGAGTTTGTAGATCGGTT 549
Db 185 GAGCTGTTGTTGGAATTTTGTGTTTCTCTTTTGTGAGAGTTTGTAGATCGGTT 244
QY 550 AGAGTTCGGCTTTACCTGGGAAGAGCAAACTTGGTGCACGCTTTCTGGACTAATT 609
Db 245 AGAGTTCGGCTTTATGTTGGACGAGAGAAAAGCTTGTCTATGCTTTCTGGACTAATT 304
QY 610 GAAGAAAATGTAACCTCTGAAAATTTAGCCCTTATCAAAAAGATGATGAGGCTAT 669
Db 305 GAAGAAAATGTAACCTCTGAAAATTTAGCCCTTGTCAAAAAGATGATGAGGCTAT 364
QY 670 GAAGTAGAGTCACTTTAGAGGATGCCAGCTTTGAGAGGAGGAGCAGAGAGACGAGA 729
Db 365 GAAGTAGAGTCACTTTAAGGATGCCAGCTTTGAGAGGAGGAGCAGAGAGACGAGA 421
QY 730 AGTTGGAGCAACCTGTGAAAAGTGAACAGGTCCAAATCTGAACTTTGAGAGTGAATC 789
Db 422 AGTTGGAGCAACCTGTGAAAAGTGAACAGGTCCAAATCTGAACTTTGAGAGTGAATC 481
QY 790 CTCTGCTAGAAAAGACTTTAAAACAAGAGAAATCTAAACATTTCTCAACAGATGAATTG 849
Db 482 CTCTGCTAGAAAAGAGATTAAAAGAGAGAAATCTAAACATTTCTCAACAGATGAATTG 541
QY 850 ATGGGGATATTTCAAAAAGTATACAGTCTCTAGAGATGAGTCAAAATCCCTCAATCA 909
Db 542 ATGGGGATATTTCAAAAAGTATACAGTCTCTAGAGATGAGTCAAAATCCCTCAATCA 601
QY 910 CAATAGCTGAGCCAAATCATCTGCAAGACATTTAAATGAGTGAAGAGCAGCGGCT 969
Db 602 CAAGTAGCTGAGCCAAATGACCTTCCAGATATTTCAATGAATGAAGAGCAGCTGAAG 661
QY 970 ATAGCAATAAAAGATGCTTTGAATGAAAATTTCTCAACTTCAGACAGCCATAAAACAGCTT 1029
Db 662 ATAGCAATAAAAGATGCTTTGAATGAAAATTTCTCAACTTCAGACAGCCATAAAACAGCTT 721
QY 1030 TTTCAGCAAGAGCTGAGTATGGAAGAGAGAGTGAAGTGAATTAATAACAGAAATA 1089
Db 722 TT---GCAAGAGCTGAGATGGAAGAGAGCAAGTGAAGTGAATTAATAACAGAAATA 778
QY 1090 ACATTTGAAGACTCCAAAGTACACGAGAACCAAGTTCTGAATGATAAAGAAAATCACATC 1149
Db 779 ACATTTGAAGACTCCAAAGTACATGCAAGAACCAAGTTCTGAATGATAAAGAACACATC 838
QY 1150 AAGACCTGACTGGACACTTGCATGATGAAGATCAGGCTGCTGTGCTTGAAGAGAC 1209
Db 839 AAGACTCTGACTGAACGCTTGTAAAGATGAAGATTTGGGCTGCTATGCTTGGAGAGAC 898
QY 1210 ACAACGGATGATGATAAAGCTGGAATGAAAGTGAACAGTGAATCGGAAAATGGTCTTAC 1269
Db 899 ATAAACGGATGATGATAAAGCTGGAATGAAAGTGAACAGTGAATCGGAAAATGGTCTTAC 958
QY 1270 TTAGATAATCTCCAAAAGAGGCTTTGAGAGAACTGATTCATGCTGCTGAAGTTAAATGCT 1329
Db 959 TTAGATAATCTCCAAAAGAGGCTTTGAGAGAACTGATTCATGCTGCTGAAGTTAAATGCT 1018
QY 1330 TCTTTAAAACCTTTAGAGAGAGAGAAACCAAAATTTATATTGCTGCTGAAGTTGAT 1389
Db 1019 TCTTTAAAACCTTTAGAGAGAGAGAAACCAAAATTTATATTGCTGCTGAAGTTGAT 1078
QY 1390 AAAACAAAAGAGAGCTTTACAGAGCATATTAATAATCTTCAGACTCAACAGACTCTTTG 1449
Db 1079 AAAACAAAAGAGAGCTTTACAGAGCATATTAATAATCTTCAGACTCAACAGACTCTTTG 1138
QY 1450 CAGTCAGAAAACACATTTTGAATGAGATCAGAGCTTCAACAGAGAACTTTAAAGTA 1509
Db 1139 CAGTCAGAAAACACATTTTGAATGAGATCAGAGCTTCAACAGAGAACTTTAAAGTA 1198

QY 1510 ATGACTGAATTTATCAAGAAAATGAAATGAAATCCACAGGAAATTAACAGTAGAGAA 1569
Db 1199 ATGACTGAATTTATCAAGAAAATGAAATGAAATCCACAGGAAATTAACAGTAGAGAA 1258
QY 1570 AATTTCGGTTAGAGAGAGAGAGAACTTTCTAAGTAGATGAAAGATCAGCCATGCC 1629
Db 1259 AATTTCGGTTAGAGAGAGAGAGAACTTTCTAAGTAGATGAAAGATCAGCCATGCC 1318
QY 1630 ACTGAAGAGCTGAGAGACCTATAGAAAAGCAGGCAAAAGATCTTGAAGAAATTTGGAGAGA 1689
Db 1319 ACTGAAGAGCTGAGAGACCTATAGAAAAGCAGGCAAAAGATCTTGAAGAAATTTGGAGAGA 1378
QY 1690 ACTATTCAATCTTATCAAGGCGAGATTTTCCATGAGAAAAGACCATGATTAATTCG 1749
Db 1379 ACTATTCAATCTTATCAAGGCGAGATTTTCCATGAGAAAAGACCATGATTAATTCG 1438
QY 1750 TTGGCAGCTCGGAATGCTGAAAAGAAACCTCAATGATTTAAGGAAAAGAAAATGCTCAAC 1809
Db 1439 TTGGCAGCTCGGAATGCTGAAAAGAAACCTCAATGATTTAAGGAAAAGAAAATGCTCAAC 1498
QY 1810 AGACAAAATTAATCTGAAAAGCAGAGCTTAAATTTGAATTTTGAAGAAAGATCTTATGCA 1869
Db 1499 AGACAAAATTAATCTGAAAAGCAGAGCTTAAATTTGAATTTTGAAGAAAGATCTTATGCA 1558
QY 1870 CTGATGTTCCAAATACAGCATTTGGCAGAGAGCATTCGCCATATGCTCCCTCACCATTTG 1929
Db 1559 CTGATGTTCCAAATACAGCATTTGGC----- 1585
QY 1930 GGTTCGCTTCATCTGAAAAGCAGAGCTTTCTCTCTCTCCAACTTTGTTGGAGGTCGA 1989
Db 1586 ----- 1585
QY 1990 CTCAGACTCTCACCTTTTCTCCAGGGGAGGAGGAGGCTCAGAGGCCCCAGGGAAT 2049
Db 1586 -----AGAGGCTCAGAGGCCCCAGGGAAT 1609
QY 2050 CCTCTGACCATCAGATTTACCAATGAAGAGAGAAATCAAGCTGTGATAGTTAAACCGAT 2109
Db 1610 CCTCTGACCATCAGATTTACCAATGAAGAGAGAAATCAAGCTGTGATAGTTAAACCGAT 1669
QY 2110 CCTCATAGGCTCCCTCTGACACTGGCTCTCTGACCTCCATGGGACAGGACCGTAGG 2169
Db 1670 CCTCATAGGCTCCCTCTGACACTGGCTCTCTGACCTCCATGGGACAGGACCGTAGG 1729
QY 2170 ATGATGTTTCTCCGCGCAGGACCAATCATATCTCTGATTCAGCCCTTCTCCACAAAGGCAA 2229
Db 1730 ATGATGTTTCTCCGCGCAGGACCAATCATATCTCTGATTCAGCCCTTCTCCACAAAGGCAA 1789
QY 2230 GACAGATTTTGTCTAATCTGCTGAGACTGTCTGGACAGGAGAACTCAGAGTTTAAAT 2289
Db 1790 GACAGATTTTGTCTAATCTGCTGAGACTGTCTGGACAGGAGAACTCAGAGTTTAAAT 1849
QY 2290 ATGCTTTCTTTGGATAAAATGGAATGGGTCAATGCTTCAGAAATGGAATCCAGTAGAAAT 2349
Db 1850 ATGCTTTCTTTGGATAAAATGGAATGGGTCAATGCTTCAGAAATGGAATCCAGTAGAAAT 1909
QY 2350 GATACCAAGAGATCTTGTGAATTTAATGCTGCTGATTCATCTCTCCCTGCTGAAAT 2409
Db 1910 GATACCAAGAGATCTTGTGAATTTAATGCTGCTGATTCATCTCTCCCTGCTGAAAT 1969
QY 2410 GAAGCCACTGCGCTTGTTCCTCCACCTTCTGCTCCAAATCAGAGGTCCTTGTGTTT 2469
Db 1970 GAAGCCACTGCGCTTGTTCCTCCACCTTCTGCTCCAAATCAGAGGTCCTTGTGTTT 2029
QY 2470 CCAGTGATGCAAGAGGCCCATTTCTTGAAGAGGAGCCTCTTTCGCCACCTCTCCCA 2529
Db 2030 CCAGTGATGCAAGAGGCCCATTTCTTGAAGAGGAGCCTCTTTCGCCACCTCTCCCA 2089
QY 2530 GGAGCCATGTTGGAGCTTCTCGAGATTTATTTCCACCAAGGAGATTTCCAGGTCACCA 2589
Db 2090 GGAGCCATGTTGGAGCTTCTCGAGATTTATTTCCACCAAGGAGATTTCCAGGTCACCA 2149

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2004, 06:17:26 ; Search time 9919 Seconds
(without alignments)
11223.801 Million cell updates/sec

Title: US-09-744-314-14

Perfect score: 3164

Sequence: 1 aggttcaatccatgaagaag.....agtgatttaaaaaaaaaa 3164

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 37577330 seqs, 17593059518 residues

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents NA Main.*

- 1: /cgn2_6/ptodata/2/pna/US091A COMB.seq.*
- 2: /cgn2_6/ptodata/2/pna/US091B COMB.seq.*
- 3: /cgn2_6/ptodata/2/pna/US091C COMB.seq.*
- 4: /cgn2_6/ptodata/2/pna/US091D COMB.seq.*
- 5: /cgn2_6/ptodata/2/pna/US091E COMB.seq.*
- 6: /cgn2_6/ptodata/2/pna/US091F COMB.seq.*
- 7: /cgn2_6/ptodata/2/pna/US091G COMB.seq.*
- 8: /cgn2_6/ptodata/2/pna/US091H COMB.seq.*
- 9: /cgn2_6/ptodata/2/pna/US091I COMB.seq.*
- 10: /cgn2_6/ptodata/2/pna/US091J COMB.seq.*
- 11: /cgn2_6/ptodata/2/pna/US091K COMB.seq.*
- 12: /cgn2_6/ptodata/2/pna/US091L COMB.seq.*
- 13: /cgn2_6/ptodata/2/pna/US091M COMB.seq.*
- 14: /cgn2_6/ptodata/2/pna/US091N COMB.seq.*
- 15: /cgn2_6/ptodata/2/pna/US091O COMB.seq.*
- 16: /cgn2_6/ptodata/2/pna/US091P COMB.seq.*
- 17: /cgn2_6/ptodata/2/pna/US091Q COMB.seq.*
- 18: /cgn2_6/ptodata/2/pna/US091R COMB.seq.*
- 19: /cgn2_6/ptodata/2/pna/US091S COMB.seq.*
- 20: /cgn2_6/ptodata/2/pna/US091T COMB.seq.*
- 21: /cgn2_6/ptodata/2/pna/US091U COMB.seq.*
- 22: /cgn2_6/ptodata/2/pna/US091V COMB.seq.*
- 23: /cgn2_6/ptodata/2/pna/US091W COMB.seq.*
- 24: /cgn2_6/ptodata/2/pna/US091X COMB.seq.*
- 25: /cgn2_6/ptodata/2/pna/US091Y COMB.seq.*
- 26: /cgn2_6/ptodata/2/pna/US091Z COMB.seq.*
- 27: /cgn2_6/ptodata/2/pna/US092A COMB.seq.*
- 28: /cgn2_6/ptodata/2/pna/US092B COMB.seq.*
- 29: /cgn2_6/ptodata/2/pna/US092C COMB.seq.*
- 30: /cgn2_6/ptodata/2/pna/US092D COMB.seq.*
- 31: /cgn2_6/ptodata/2/pna/US092E COMB.seq.*
- 32: /cgn2_6/ptodata/2/pna/US092F COMB.seq.*
- 33: /cgn2_6/ptodata/2/pna/US092G COMB.seq.*
- 34: /cgn2_6/ptodata/2/pna/US092H COMB.seq.*
- 35: /cgn2_6/ptodata/2/pna/US092I COMB.seq.*
- 36: /cgn2_6/ptodata/2/pna/US092J COMB.seq.*
- 37: /cgn2_6/ptodata/2/pna/US092K COMB.seq.*
- 38: /cgn2_6/ptodata/2/pna/US092L COMB.seq.*
- 39: /cgn2_6/ptodata/2/pna/US092M COMB.seq.*
- 40: /cgn2_6/ptodata/2/pna/US092N COMB.seq.*
- 41: /cgn2_6/ptodata/2/pna/US092O COMB.seq.*
- 42: /cgn2_6/ptodata/2/pna/US092P COMB.seq.*
- 43: /cgn2_6/ptodata/2/pna/US092Q COMB.seq.*

- 44: /cgn2_6/ptodata/2/pna/US100B COMB.seq.*
- 45: /cgn2_6/ptodata/2/pna/US101A COMB.seq.*
- 46: /cgn2_6/ptodata/2/pna/US101B COMB.seq.*
- 47: /cgn2_6/ptodata/2/pna/US102A COMB.seq.*
- 48: /cgn2_6/ptodata/2/pna/US102B COMB.seq.*
- 49: /cgn2_6/ptodata/2/pna/US103A COMB.seq.*
- 50: /cgn2_6/ptodata/2/pna/US103B COMB.seq.*
- 51: /cgn2_6/ptodata/2/pna/US104A COMB.seq.*
- 52: /cgn2_6/ptodata/2/pna/US104B COMB.seq.*
- 53: /cgn2_6/ptodata/2/pna/US105 COMB.seq.*
- 54: /cgn2_6/ptodata/2/pna/US107A COMB.seq.*
- 55: /cgn2_6/ptodata/2/pna/US107B COMB.seq.*
- 56: /cgn2_6/ptodata/2/pna/US6000 COMB.seq.*
- 57: /cgn2_6/ptodata/2/pna/US6001 COMB.seq.*
- 58: /cgn2_6/ptodata/2/pna/US6002 COMB.seq.*
- 59: /cgn2_6/ptodata/2/pna/US6003 COMB.seq.*
- 60: /cgn2_6/ptodata/2/pna/US6004 COMB.seq.*
- 61: /cgn2_6/ptodata/2/pna/US6005 COMB.seq.*
- 62: /cgn2_6/ptodata/2/pna/US6006 COMB.seq.*
- 63: /cgn2_6/ptodata/2/pna/US6007 COMB.seq.*
- 64: /cgn2_6/ptodata/2/pna/US6008 COMB.seq.*
- 65: /cgn2_6/ptodata/2/pna/US6009 COMB.seq.*
- 66: /cgn2_6/ptodata/2/pna/US6010 COMB.seq.*
- 67: /cgn2_6/ptodata/2/pna/US6011 COMB.seq.*
- 68: /cgn2_6/ptodata/2/pna/US6012 COMB.seq.*
- 69: /cgn2_6/ptodata/2/pna/US6013 COMB.seq.*
- 70: /cgn2_6/ptodata/2/pna/US6014 COMB.seq.*
- 71: /cgn2_6/ptodata/2/pna/US6015 COMB.seq.*
- 72: /cgn2_6/ptodata/2/pna/US6016 COMB.seq.*
- 73: /cgn2_6/ptodata/2/pna/US6017 COMB.seq.*
- 74: /cgn2_6/ptodata/2/pna/US6018 COMB.seq.*
- 75: /cgn2_6/ptodata/2/pna/US6019 COMB.seq.*
- 76: /cgn2_6/ptodata/2/pna/US6020 COMB.seq.*
- 77: /cgn2_6/ptodata/2/pna/US6021 COMB.seq.*
- 78: /cgn2_6/ptodata/2/pna/US6022 COMB.seq.*
- 79: /cgn2_6/ptodata/2/pna/US6023A COMB.seq.*
- 80: /cgn2_6/ptodata/2/pna/US6023B COMB.seq.*
- 81: /cgn2_6/ptodata/2/pna/US6024 COMB.seq.*
- 82: /cgn2_6/ptodata/2/pna/US6025 COMB.seq.*
- 83: /cgn2_6/ptodata/2/pna/US6026 COMB.seq.*
- 84: /cgn2_6/ptodata/2/pna/US6027 COMB.seq.*
- 85: /cgn2_6/ptodata/2/pna/US6028 COMB.seq.*
- 86: /cgn2_6/ptodata/2/pna/US6029 COMB.seq.*
- 87: /cgn2_6/ptodata/2/pna/US6030 COMB.seq.*
- 88: /cgn2_6/ptodata/2/pna/US6031 COMB.seq.*
- 89: /cgn2_6/ptodata/2/pna/US6032 COMB.seq.*
- 90: /cgn2_6/ptodata/2/pna/US6033 COMB.seq.*
- 91: /cgn2_6/ptodata/2/pna/US6034 COMB.seq.*
- 92: /cgn2_6/ptodata/2/pna/US6035 COMB.seq.*
- 93: /cgn2_6/ptodata/2/pna/US6036 COMB.seq.*
- 94: /cgn2_6/ptodata/2/pna/US6037 COMB.seq.*
- 95: /cgn2_6/ptodata/2/pna/US6038 COMB.seq.*
- 96: /cgn2_6/ptodata/2/pna/US6039 COMB.seq.*
- 97: /cgn2_6/ptodata/2/pna/US6040 COMB.seq.*
- 98: /cgn2_6/ptodata/2/pna/US6041 COMB.seq.*
- 99: /cgn2_6/ptodata/2/pna/US6042 COMB.seq.*
- 100: /cgn2_6/ptodata/2/pna/US6043 COMB.seq.*
- 101: /cgn2_6/ptodata/2/pna/US6044 COMB.seq.*
- 102: /cgn2_6/ptodata/2/pna/US6045 COMB.seq.*
- 103: /cgn2_6/ptodata/2/pna/US6046 COMB.seq.*
- 104: /cgn2_6/ptodata/2/pna/US6047 COMB.seq.*
- 105: /cgn2_6/ptodata/2/pna/US6048 COMB.seq.*
- 106: /cgn2_6/ptodata/2/pna/US6049 COMB.seq.*
- 107: /cgn2_6/ptodata/2/pna/US6050 COMB.seq.*
- 108: /cgn2_6/ptodata/2/pna/US6051 COMB.seq.*
- 109: /cgn2_6/ptodata/2/pna/US6052 COMB.seq.*
- 110: /cgn2_6/ptodata/2/pna/US6053 COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

;; PRIOR FILING DATE: 1998-07-31; 1998-07-31; 1998-08-04; 1998-08-04

17

;; NUMBER OF SEQ ID NOS:

;; SOFTWARE: PERL Program

;; SEQ ID NO 14

;; LENGTH: 3164

;; TYPE: DNA

;; ORGANISM: Homo sapiens

;; FEATURE:

;; NAME/KEY: misc feature

;; OTHER INFORMATION: Incyte Clone No: 3768043

PCT-US99-17167-14

Query Match 100.0%; Score 3164; DB 1; Length 3164;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTTTAAATCCATGAAGACACACAAATTTTAAAGTGATTTCACCAAAAAATAAGCTT 60
 DB 1 AGTTTAAATCCATGAAGACACACAAATTTTAAAGTGATTTCACCAAAAAATAAGCTT 60
 QY 61 CAAATATGTGATGTGAAACTCCAGAACTAGGGGGGCGGGCTCAGACAGCGCTGC 120
 DB 61 CAAATATGTGATGTGAAACTCCAGAACTAGGGGGGCGGGCTCAGACAGCGCTGC 120
 QY 121 CTGAGGATGTAAGTGTAAACAAGAGGCGCAGGGAGGTGTGGGGGACACATGGGCGCTG 180
 DB 121 CTGAGGATGTAAGTGTAAACAAGAGGCGCAGGGAGGTGTGGGGGACACATGGGCGCTG 180
 QY 181 TGAGGCGTGTGGTCCCGGCTTCCAGCTCCCGGCGAGCCCGCTCCACAGTGTCCG 240
 DB 181 TGAGGCGTGTGGTCCCGGCTTCCAGCTCCCGGCGAGCCCGCTCCACAGTGTCCG 240
 QY 241 CTCGGTGTGGTGTACGTGCGGATTCGGGTTCAGACCAAGCGTCCGTGTCTCCACC 300
 DB 241 CTCGGTGTGGTGTACGTGCGGATTCGGGTTCAGACCAAGCGTCCGTGTCTCCACC 300
 QY 301 GCTTGTGTGGCAGTGTATCTGGGTGACCGCCAGAGAGCTTCAGCGTATGGAGGAG 360
 DB 301 GCTTGTGTGGCAGTGTATCTGGGTGACCGCCAGAGAGCTTCAGCGTATGGAGGAG 360
 QY 361 CCTGGTGTACCCCTCAGCCCTACCTGGGCTGTCTGGAGGAGTACGACAGTGTG 420
 DB 361 CCTGGTGTACCCCTCAGCCCTACCTGGGCTGTCTGGAGGAGTACGACAGTGTG 420
 QY 421 GCACCACTACCTGAGATGTAGACACAGATGAGAACTCTTATGGTTTCCATCGAAGCTG 480
 DB 421 GCACCACTACCTGAGATGTAGACACAGATGAGAACTCTTATGGTTTCCATCGAAGCTG 480
 QY 481 GTGATGTGACAGCTGTATTGGATTTTTTGTGTTCTCCCTTTTTTGTGGAGAGTTT 540
 DB 481 GTGATGTGACAGCTGTATTGGATTTTTTGTGTTCTCCCTTTTTTGTGGAGAGTTT 540
 QY 541 AGATCGGTTAGGAGTCCGCTTTTACGTGGGAGAGAGCAAAAACCTGGTGCACCGTTTCT 600
 DB 541 AGATCGGTTAGGAGTCCGCTTTTACGTGGGAGAGAGCAAAAACCTGGTGCACCGTTTCT 600
 QY 601 GGATTAATGAGAGAAATCTTAACCTACTTTGAAAAATTTAGCCTTATTTCAAAAGAGTAT 660
 DB 601 GGATTAATGAGAGAAATCTTAACCTACTTTGAAAAATTTAGCCTTATTTCAAAAGAGTAT 660
 QY 661 GAAGGCTATGAAGTAGAGTATCATCTTTAGAGGATGCCAGCTTTGAGAGGAGGAGCAGAA 720
 DB 661 GAAGGCTATGAAGTAGAGTATCATCTTTAGAGGATGCCAGCTTTGAGAGGAGGAGCAGAA 720
 QY 721 GAAGCAGAGAGTTGGAGGCAACCTGTGAAAGCTGAACAGGTCCTAATCTGACCTTGG 780
 DB 721 GAAGCAGAGAGTTGGAGGCAACCTGTGAAAGCTGAACAGGTCCTAATCTGACCTTGG 780
 QY 781 GATGAAATCCCTCTCTGTAGAAAAAGACTTAAAAACAAGAGAAATCTAAACATTTCTCAACAA 840
 DB 781 GATGAAATCCCTCTCTGTAGAAAAAGACTTAAAAACAAGAGAAATCTAAACATTTCTCAACAA 840
 QY 841 GATGAAATGATGGCGGATATTTTCAAAAGATATACAGTCTCTAGAGATGAGTCAAAATCC 900

Description

1 3164 100.0 3164 1 PCT-US99-17167-14 Sequence 14, Appl
 2 3164 100.0 3164 15 US-09-127-665-6 Sequence 6, Appl
 3 3164 100.0 3164 31 US-09-744-314-14 Sequence 14, Appl
 4 3164 100.0 3164 55 US-10-753-200-320 Sequence 320, App
 5 3164 100.0 3164 81 US-60-240-409-320 Sequence 320, App
 6 3159.2 99.8 3164 89 US-60-324-185-1645 Sequence 1645, App
 7 3153 99.7 3165 73 US-60-172-360-25590 Sequence 25590, A
 8 2880.4 91.0 3108 1 PCT-US02-25465-51 Sequence 51, Appl
 9 2836.4 89.6 12244 107 US-60-500-315-11859 Sequence 11859, A
 10 2780.2 87.9 3659 39 US-09-948-941-107 Sequence 107, App
 11 2780.2 87.9 3659 46 US-10-170-235-34751 Sequence 34751, A
 12 2773.8 87.7 3310 32 US-09-770-175-5675 Sequence 6675, App
 13 2764.2 87.4 3676 39 US-09-948-941-74 Sequence 74, Appl
 14 2764.2 87.4 3676 106 US-60-490-890-33 Sequence 33, Appl
 15 2748.6 86.9 3897 27 US-09-338-425-2509 Sequence 2509, App
 16 2748.6 86.9 3897 27 US-09-644-867-7218 Sequence 7218, App
 17 2748.6 86.9 3897 27 US-09-652-128-9585 Sequence 9585, App
 18 2748.6 86.9 3897 27 US-09-652-127-9671 Sequence 9671, App
 19 2748.6 86.9 3897 27 US-09-652-128-9776 Sequence 9776, App
 20 2748.6 86.9 3897 27 US-09-652-914-9084 Sequence 9084, App
 21 2748.6 86.9 3897 29 US-09-698-010-13064 Sequence 13064, A
 22 2748.6 86.9 3897 29 US-09-698-012-7978 Sequence 7978, App
 23 2748.6 86.9 3897 29 US-09-698-013-6459 Sequence 6459, App
 24 2748.6 86.9 3897 29 US-09-698-999-9200 Sequence 6200, App
 25 2748.6 86.9 3897 31 US-09-726-170-2155 Sequence 2155, App
 26 2748.6 86.9 3897 31 US-09-726-788-6235 Sequence 6235, App
 27 2748.6 86.9 3897 31 US-09-726-790-2509 Sequence 2509, App
 28 2567.8 81.2 3681 1 PCT-US01-08631-5271 Sequence 5271, App
 29 2512.2 79.5 2928 1 PCT-US02-25465-38 Sequence 38, Appl
 30 2501.2 79.4 3796 46 US-10-170-235-34682 Sequence 34682, A
 31 2501.4 79.1 3915 1 PCT-US01-08631-20249 Sequence 20249, A
 32 2495 78.9 20719 77 US-60-212-664-73 Sequence 73, Appl
 33 2489.8 78.7 14647 103 US-60-461-762-3548 Sequence 3548, App
 34 2489.8 78.7 14647 104 US-60-470-166-9150 Sequence 9150, App
 35 2489.8 78.7 60686 103 US-60-461-762-3357 Sequence 3357, App
 36 2489.8 78.7 60686 104 US-60-470-166-8696 Sequence 8696, App
 37 2483.4 78.5 3296 47 US-10-236-177-179 Sequence 179, App
 38 2447.6 77.4 2884 84 US-60-278-232-701 Sequence 701, App
 39 2416.2 76.4 2770 55 US-10-753-200-322 Sequence 322, App
 40 2416.2 76.4 2770 81 US-60-240-409-322 Sequence 322, App
 41 2416.2 76.4 3456 73 US-60-172-360-21252 Sequence 21252, A
 42 2414.6 76.3 2814 36 US-09-880-107-3419 Sequence 3419, App
 43 2414.6 76.3 2814 41 US-09-967-768A-191 Sequence 191, App
 44 2414.6 76.3 2814 41 US-09-967-768C-191 Sequence 191, App
 45 2414.6 76.3 2814 44 US-10-081-056-335 Sequence 335, App

ALIGNMENTS

RESULT 1

PCT-US99-17167-14

;; Sequence 14, Application PC/TUS9917167

;; GENERAL INFORMATION:

;; APPLICANT: INCYTE PHARMACEUTICALS, INC.

;; APPLICANT: BANDMAN, Olga

;; APPLICANT: TANG, Y. Tom

;; APPLICANT: YUE, Henry

;; APPLICANT: CORLEY, Neil C.

;; APPLICANT: GUEGLER, Karl J.

;; APPLICANT: AZIMZAI, Yalda

;; APPLICANT: PATTERSON, Chandra

;; APPLICANT: LAL, Preeti

;; APPLICANT: BAUGHN, Mariah R.

;; TITLE OF INVENTION: HUMAN CYTOSKELETAL PROTEINS

;; FILE REFERENCE: PF-0568 PCT

;; CURRENT APPLICATION NUMBER: PCT/US99/17167

;; PRIOR FILING DATE: 1999-09-21

;; CURRENT APPLICATION NUMBER: 09/127,665; unassigned; 09/128,850; unassigned

841 GATGAATGATGGCGATATTTCCAAAAGATATACAGTCTCTAGAGATGAGTCAAAATCC 900
901 CTCAAATCAAAATAGCTGAGAGCCAAATATCATCTGCAAGACATTTAAATAGTGAAGAA 960
901 CTCAAATCAAAATAGCTGAGAGCCAAATATCATCTGCAAGACATTTAAATAGTGAAGAA 960
961 CGACGGGCTATAGCAATAAAGATGCTTTGAATGAAAATTTCTCAACTTCAGACAAGCCAT 1020
961 CGACGGGCTATAGCAATAAAGATGCTTTGAATGAAAATTTCTCAACTTCAGACAAGCCAT 1020
1021 AAACAGCTTTTTCAGCAAGAGCTGAGTATGAAAGGAGAGTGAAGTAAATTAATAA 1080
1021 AAACAGCTTTTTCAGCAAGAGCTGAGTATGAAAGGAGAGTGAAGTAAATTAATAA 1080
1081 CAGAAAATAAATTTCAAGAGCTTCAAGTACACGACAGAAACAAAGTTCTGAAATCAATAAGAA 1140
1081 CAGAAAATAAATTTCAAGAGCTTCAAGTACACGACAGAAACAAAGTTCTGAAATCAATAAGAA 1140
1141 AATCACATCAAGACCTGACTGTGACACTTGCCTCAATGATGAAAGATCAGGCTGCTGTGCTT 1200
1141 AATCACATCAAGACCTGACTGTGACACTTGCCTCAATGATGAAAGATCAGGCTGCTGTGCTT 1200
1201 GAAGAGACACACAGCGATGATGATACCTGGATTTAGAGTGAACAGTGAATCGGAAT 1260
1201 GAAGAGACACACAGCGATGATGATACCTGGATTTAGAGTGAACAGTGAATCGGAAT 1260
1261 GGTGCTTACTTAGATAATCTTCAAAAGAGCTTTTGAAGAACTGATTCATGCTGCTAAG 1320
1261 GGTGCTTACTTAGATAATCTTCAAAAGAGCTTTTGAAGAACTGATTCATGCTGCTAAG 1320
1321 TTAATGCTTCTTTAAACCTTTAGAGGAGAAAGAAACCAATTTATATTCAGTGTGCT 1380
1321 TTAATGCTTCTTTAAACCTTTAGAGGAGAAAGAAACCAATTTATATTCAGTGTGCT 1380
1381 GAAGTTGATAAAACAAAGGAGAGCTTTACAGAGCATATTAATAATCTTCAGACTCAACAA 1440
1381 GAAGTTGATAAAACAAAGGAGAGCTTTACAGAGCATATTAATAATCTTCAGACTCAACAA 1440
1441 GCATCTTGAGCTGAGAGAAACACACATTTTGAATAAGAGATCAGAGCTTCAACAGAAA 1500
1441 GCATCTTGAGCTGAGAGAAACACACATTTTGAATAAGAGATCAGAGCTTCAACAGAAA 1500
1501 CTTAAAGTAACTGATGATATATCAAGAAATGAAATGAACTCCACAGGAAATTAACA 1560
1501 CTTAAAGTAACTGATGATATATCAAGAAATGAAATGAACTCCACAGGAAATTAACA 1560
1561 GTAGAGAAATATTCGGTTAGAGAAAGAGAGAACTTTCTAAAGTAGATGAAAGATC 1620
1561 GTAGAGAAATATTCGGTTAGAGAAAGAGAGAACTTTCTAAAGTAGATGAAAGATC 1620
1621 AGCCATGCCACTGAGAGCTGAGACCTATAGAAAGCGAGCCAAAGATCTTGAAGAGAA 1680
1621 AGCCATGCCACTGAGAGCTGAGACCTATAGAAAGCGAGCCAAAGATCTTGAAGAGAA 1680
1681 TTGAGAGAACTATTCATCTTCAAGGCGAGATTTTCCCATGAGAAAAGCACAT 1740
1681 TTGAGAGAACTATTCATCTTCAAGGCGAGATTTTCCCATGAGAAAAGCACAT 1740
1741 GATAATTTGGTGGCAGCTCGGAATGCTGAAGAAACCTCAATGATTTAGGAAGAAAT 1800
1741 GATAATTTGGTGGCAGCTCGGAATGCTGAAGAAACCTCAATGATTTAGGAAGAAAT 1800
1801 GCTCACAAACAGACAAAATTAACCTGAAACAGAGCTTAAATTTGAACCTTTTGAAGAAAGAT 1860
1801 GCTCACAAACAGACAAAATTAACCTGAAACAGAGCTTAAATTTGAACCTTTTGAAGAAAGAT 1860
1861 CCCTATGCACTCGATGTTCCAAATACAGCATTTGGCAGAGAGCATTTCCCATATGCTCCC 1920
1861 CCCTATGCACTCGATGTTCCAAATACAGCATTTGGCAGAGAGCATTTCCCATATGCTCCC 1920
1921 TCACATTTGGTGGCCTTCACTGAAACCAAGAGCTTTTCTCTCTCCCAACTTTTGTG 1980

1921 TCACATTTGGTGGCCTTCACTGAAACCAAGAGCTTTTCTCTCTCCCAACTTTTGTG 1980
1981 GAGGCTCACTCAGACTCTCAGCTTTGCTTCCAGGGGAGAGAGAGAGCTCAGGAGC 2040
1981 GAGGCTCACTCAGACTCTCAGCTTTGCTTCCAGGGGAGAGAGAGAGCTCAGGAGC 2040
2041 CCAGGGAATCTCTCGGACCATCAGATTACCAATGAAAGAGAGAGAAATCAAGCTGTGATAG 2100
2041 CCAGGGAATCTCTCGGACCATCAGATTACCAATGAAAGAGAGAGAAATCAAGCTGTGATAG 2100
2101 TTAACCATCTCTATAGGCTCCCTCTGACACTGGGTCTCTGTCCCTCCATCGGAGCAG 2160
2101 TTAACCATCTCTATAGGCTCCCTCTGACACTGGGTCTCTGTCCCTCCATCGGAGCAG 2160
2161 GACCTAGGATGATGTTTCTCTCGCCAGGACAAATCATATCTCTGATTGAGCCCTTCTCTCA 2220
2161 GACCTAGGATGATGTTTCTCTCGCCAGGACAAATCATATCTCTGATTGAGCCCTTCTCTCA 2220
2221 CAAAGGCAAGACAGATTTGTTCTAATCTGTGTAGACTGTCTGACCCAGCAGAGACTCAGA 2280
2221 CAAAGGCAAGACAGATTTGTTCTAATCTGTGTAGACTGTCTGACCCAGCAGAGACTCAGA 2280
2281 AGTTTAAATGCTTCTTTGGATAAATGATGGGTCAATGCTTCCAGAAATGGAATCC 2340
2281 AGTTTAAATGCTTCTTTGGATAAATGATGGGTCAATGCTTCCAGAAATGGAATCC 2340
2341 AGTGAATAATGATCAAAAGAGATCTTTGTTAAATTTAAATGTGCTGATTCATCTCTCCT 2400
2341 AGTGAATAATGATCAAAAGAGATCTTTGTTAAATTTAAATGTGCTGATTCATCTCTCCT 2400
2401 GCTGAAATGAAGCCACTGGCCCTGGCTTCTCTCCACTCTTGTCTCCAAATCAGAGGT 2460
2401 GCTGAAATGAAGCCACTGGCCCTGGCTTCTCTCCACTCTTGTCTCCAAATCAGAGGT 2460
2461 CCATTTGTTTCCAGTGGATGCAAGAGGCCCCATTTCTGAGAGAGAGAGACTCTCTTCCC 2520
2461 CCATTTGTTTCCAGTGGATGCAAGAGGCCCCATTTCTGAGAGAGAGAGACTCTCTTCCC 2520
2521 CCTCTCCAGAGAGCCATGTTTGGAGCTTCTCGAGATTTATTTCCACCAAGAGATTTCCCA 2580
2521 CCTCTCCAGAGAGCCATGTTTGGAGCTTCTCGAGATTTATTTCCACCAAGAGATTTCCCA 2580
2581 GGTCCACACCTGCTCCATTTGCAATGAGAAATGCTCTATCCACCGAGGGGTTTCTCTCT 2640
2581 GGTCCACACCTGCTCCATTTGCAATGAGAAATGCTCTATCCACCGAGGGGTTTCTCTCT 2640
2641 TACCTTCCCCCAGAGCCCTGGATTTTCCCCCAGCCCACTTCTGAGGTAGAGTGAG 2700
2641 TACCTTCCCCCAGAGCCCTGGATTTTCCCCCAGCCCACTTCTGAGGTAGAGTGAG 2700
2701 TTTCCCTCAGGTTTGAATTTCCACTTTCAAAATGAGGCTGCTACTGAAACATCCAGAACCA 2760
2701 TTTCCCTCAGGTTTGAATTTCCACTTTCAAAATGAGGCTGCTACTGAAACATCCAGAACCA 2760
2761 CAGAAACCTGACAAATATTTTGTCTCTTCAAAAGTAAATTTGACTGATCTCATTTTCA 2820
2761 CAGAAACCTGACAAATATTTTGTCTCTTCAAAAGTAAATTTTGAATCTCATTTTCA 2820
2821 GTTTAAAGTAACTGCTGTTACTTAAAGTGAATACATTTTGTCTCAAAATTTGAAGCTTAA 2880
2821 GTTTAAAGTAACTGCTGTTACTTAAAGTGAATACATTTTGTCTCAAAATTTGAAGCTTAA 2880
2881 ATTAAATCTCAGGATAGTATTTTGTAAATGAAAGATGATTTAAATATGAAATCTTATGAG 2940
2881 ATTAAATCTCAGGATAGTATTTTGTAAATGAAAGATGATTTAAATATGAAATCTTATGAG 2940
2941 TAAATTTTCAATTTTATTTAGACGATTAATCTTATTTCAATTTGATTTAAATCTTAT 3000
2941 TAAATTTTCAATTTTATTTAGACGATTAATCTTATTTCAATTTGATTTAAATCTTAT 3000
3001 ATATAAACAAATAGTGGAGTATTTATATGATTAATCTTTTCAAGGTGGGAGGCTTTAAATTC 3060
3001 ATATAAACAAATAGTGGAGTATTTATATGATTAATCTTTTCAAGGTGGGAGGCTTTAAATTC 3060

QY 3061 TGAAGTCTGCTGCTTATCCAGAGTGTATTTACTGTGGTGTGGCAATGTGAAG 3120
Db 3061 TGAAGTCTGCTGCTTATCCAGAGTGTATTTACTGTGGTGTGGCAATGTGAAG 3120
QY 3121 TAACCTTTATGCTTAAATAAATTATAGTTGATTTAAAAA 3164
Db 3121 TAACCTTTATGCTTAAATAAATTATAGTTGATTTAAAAA 3164

RESULT 2
US-09-127-665-6
; Sequence 6, Application US/09127665
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Azimzal, Yalda
; APPLICANT: Patterson, Chandra
; TITLE OF INVENTION: HUMAN TROPOMYOSIN-LIKE TUMOR-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0568 US
; CURRENT APPLICATION NUMBER: US/09/127,665
; CURRENT FILING DATE: 1998-07-31
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 3164
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
; FEATURE:
; OTHER INFORMATION: 3758043, BRSTN0724
US-09-127-665-6

Query Match 100.0%; Score 3164; DB 15; Length 3164;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGTTTAAATCCATGAAGAGACAGCAATTTTAAAGTGTATTCAACCAAAAATAAGCTT 60
Db 1 AGGTTTAAATCCATGAAGAGACAGCAATTTTAAAGTGTATTCAACCAAAAATAAGCTT 60
QY 61 CAAAATATGTGATGTGAAAACCTGCCAGAACTAAGCGCGCGCGGTTCAGACCAGCGCTGC 120
Db 61 CAAAATATGTGATGTGAAAACCTGCCAGAACTAAGCGCGCGCGGTTCAGACCAGCGCTGC 120
QY 121 CTCAGAGTGAAGTGTAAACAAGAGGCGCAGGGAGGTGGTGGGGACAACATGGGCGTG 180
Db 121 CTCAGAGTGAAGTGTAAACAAGAGGCGCAGGGAGGTGGTGGGGACAACATGGGCGTG 180
QY 181 TGAGGCTGTGGTGGCGCGGTTCCTCCAGCTCCCGCGCAGCGCGCTCCACAGTGGTCCG 240
Db 181 TGAGGCTGTGGTGGCGCGGTTCCTCCAGCTCCCGCGCAGCGCGCTCCACAGTGGTCCG 240
QY 241 CTCGGTTGGTGTACAGTGGCGATTCGGGTTCAGACCCAGGCTGGTGTCTCCACC 300
Db 241 CTCGGTTGGTGTACAGTGGCGATTCGGGTTCAGACCCAGGCTGGTGTCTCCACC 300
QY 301 GCTGTGTGGCGCAGTGTACTGCGGTGACCGCAGAGCAGCTCGACGCTATGAGGAG 360
Db 301 GCTGTGTGGCGCAGTGTACTGCGGTGACCGCAGAGCAGCTCGACGCTATGAGGAG 360
QY 361 CTTGTGTACCCCTCAGCCCTACCTGGGCTGGTCTCGAGAGCTACGCAAGTGTG 420
Db 361 CTTGTGTACCCCTCAGCCCTACCTGGGCTGGTCTCGAGAGCTACGCAAGTGTG 420
QY 421 CGAGCACTACCTGAGAGTATGAGCAGATGAGATCCTTATGTTTCCATCGGAAGT 480
Db 421 CGAGCACTACCTGAGAGTATGAGCAGATGAGATCCTTATGTTTCCATCGGAAGT 480
QY 481 GTGGTATGTCAGCTGTATTGATTTTGTGTTCTCTCTTTTGTGGAGAGT 540
Db 481 GTGGTATGTCAGCTGTATTGATTTTGTGTTCTCTCTTTTGTGGAGAGT 540

Db 481 GTGGTATGTCAGCTGTATTGATTTTGTGTTCTCTCTTTTGTGGAGAGT 540
QY 541 AGATCGGTTAGGAGTGGCTTTTACGTGGAGAGAGCAAAAACCTTGGTGGACGCTTTCT 600
Db 541 AGATCGGTTAGGAGTGGCTTTTACGTGGAGAGAGCAAAAACCTTGGTGGACGCTTTCT 600
QY 601 GGACTAATTCAGAGAAAAATGTAACCTACTTGAATAAATTTAGCCTTATTCAAAAAGAGTAT 660
Db 601 GGACTAATTCAGAGAAAAATGTAACCTACTTGAATAAATTTAGCCTTATTCAAAAAGAGTAT 660
QY 661 GAAGGCTATGAAGTAGAGTATCTTTAGAGGATGCCAGCTTTGAGAGAGGAGGAGAGAA 720
Db 661 GAAGGCTATGAAGTAGAGTATCTTTAGAGGATGCCAGCTTTGAGAGAGGAGGAGAGAA 720
QY 721 GAAGCACCAAGTTTGGAGGCAACCTGTGAAAAGCTGAAACAGGTCCTCAATTCGAATTTGAG 780
Db 721 GAAGCACCAAGTTTGGAGGCAACCTGTGAAAAGCTGAAACAGGTCCTCAATTCGAATTTGAG 780
QY 781 GATGAATTCCTCTGTCTAGAAAAAGCTTAAACCAAGAGAAATCTAAACATTTCTCAACAA 840
Db 781 GATGAATTCCTCTGTCTAGAAAAAGCTTAAACCAAGAGAAATCTAAACATTTCTCAACAA 840
QY 841 GATGAATTCCTGTGCGGATATTTCAAAAAGTATACAGTCTCTAGAGATGAGTCAAAATCC 900
Db 841 GATGAATTCCTGTGCGGATATTTCAAAAAGTATACAGTCTCTAGAGATGAGTCAAAATCC 900
QY 901 CTCAAATCAAAATAGCTGAAGCCAAAATCATCTGCAAGACATTTAAATAGTGAAGAA 960
Db 901 CTCAAATCAAAATAGCTGAAGCCAAAATCATCTGCAAGACATTTAAATAGTGAAGAA 960
QY 961 CGACGGCTATAGCAATAAAGATGCTTTGAATGAAATTTCAACTTCAGACAAGCCAT 1020
Db 961 CGACGGCTATAGCAATAAAGATGCTTTGAATGAAATTTCAACTTCAGACAAGCCAT 1020
QY 1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGAAAAGGAGAGTGAAGTGAATTAATAA 1080
Db 1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGAAAAGGAGAGTGAAGTGAATTAATAA 1080
QY 1081 CAGAAATACATTTGAGAGCTCCAAAGTACACGAGACACAGTCTTGAATGATAAGAA 1140
Db 1081 CAGAAATACATTTGAGAGCTCCAAAGTACACGAGACACAGTCTTGAATGATAAGAA 1140
QY 1141 AATCACATCAAGACCTCTGACTGGACACTTGCCTCAATGATGAAGATCAGGCTGTGCT 1200
Db 1141 AATCACATCAAGACCTCTGACTGGACACTTGCCTCAATGATGAAGATCAGGCTGTGCT 1200
QY 1201 GAAGAGACACACAGGATGATGATTAACCTGGATTTAGAGTGAACAGTGAATCGGAAT 1260
Db 1201 GAAGAGACACACAGGATGATGATTAACCTGGATTTAGAGTGAACAGTGAATCGGAAT 1260
QY 1261 GGTGCTTACTTAGATAATCTCCAAAAGGAGCTTTGAAAAGAACTGATTCATGCTGCTAAG 1320
Db 1261 GGTGCTTACTTAGATAATCTCCAAAAGGAGCTTTGAAAAGAACTGATTCATGCTGCTAAG 1320
QY 1321 TTAATGCTTCTTTAAACCTTTAGAGGAGAGAAACCAATTTATTTATTCAGTTGTCT 1380
Db 1321 TTAATGCTTCTTTAAACCTTTAGAGGAGAGAAACCAATTTATTTATTCAGTTGTCT 1380
QY 1381 GAAGTTCATAAAACAAAGGAGAGCTTTACAGAGCATATTAAAAATCTTCAGACTCAACAA 1440
Db 1381 GAAGTTCATAAAACAAAGGAGAGCTTTACAGAGCATATTAAAAATCTTCAGACTCAACAA 1440
QY 1441 GCATCTTTGAGTCAAGAAACACACATTTTGAATAATGAGANTCAGAGCTTCACAGAAA 1500
Db 1441 GCATCTTTGAGTCAAGAAACACACATTTTGAATAATGAGANTCAGAGCTTCACAGAAA 1500
QY 1501 CTTAAAGTATGATGATTAATTAATCAAGAAATGAAATGAAATCCACAGGAAATTAACA 1560
Db 1501 CTTAAAGTATGATGATTAATTAATCAAGAAATGAAATGAAATCCACAGGAAATTAACA 1560
QY 1561 GTAGAGGAAATTTATCGGTTAGAGAAAGAGAGAAATTTCTAAAGTAGATGAAAGATC 1620
Db 1561 GTAGAGGAAATTTATCGGTTAGAGAAAGAGAGAAATTTCTAAAGTAGATGAAAGATC 1620

1621 AGCCATGCCACTGAGAGCTGGAGACCTATAGAAAGCGAGCCAAAGATCTTGAAGAAGAA 1680
1621 AGCCATGCCACTGAGAGCTGGAGACCTATAGAAAGCGAGCCAAAGATCTTGAAGAAGAA 1680
1681 TTGGAGAGAACTATTCAATCTTATCAAGGCGAGATATTTCCCATGAGAAAAAGCACAT 1740
1681 TTGGAGAGAACTATTCAATCTTATCAAGGCGAGATATTTCCCATGAGAAAAAGCACAT 1740
1741 GATAATTTGGTGGCAGCTCGGAATGCTGAAAGAAACCTCAATGATTAAAGAAAGAAAT 1800
1741 GATAATTTGGTGGCAGCTCGGAATGCTGAAAGAAACCTCAATGATTAAAGAAAGAAAT 1800
1801 GCTCACAAACAGACAAATAATTAACCTGAAACAGAGCTTAAATTTGAACTTTAGAAAAGAT 1860
1801 GCTCACAAACAGACAAATAATTAACCTGAAACAGAGCTTAAATTTGAACTTTAGAAAAGAT 1860
1861 CATTATGCACTCGATGTTCCAAATACAGCAATTTGGCAGAGAGCATTTCCCATATGTCCTC 1920
1861 CATTATGCACTCGATGTTCCAAATACAGCAATTTGGCAGAGAGCATTTCCCATATGTCCTC 1920
1921 TCACCATTTGGTGGCCTTCTCATCTGAAACAGAGCTTTCTCTCTCTCTCTCTCTCTCTCT 1980
1921 TCACCATTTGGTGGCCTTCTCATCTGAAACAGAGCTTTCTCTCTCTCTCTCTCTCTCTCT 1980
1981 GAGGGTCCACTCAGACTCTCACCCTTTGCTTCCAGGGGAGGAGAGAGGCTCACGAGGC 2040
1981 GAGGGTCCACTCAGACTCTCACCCTTTGCTTCCAGGGGAGGAGAGAGGCTCACGAGGC 2040
2041 CCAGGGAATCTCTGGACCATCAGATTACCAATGAAGAGAGAGATCAAGCTGTATAGG 2100
2041 CCAGGGAATCTCTGGACCATCAGATTACCAATGAAGAGAGAGATCAAGCTGTATAGG 2100
2101 TTAACCGATCTCATAGGGTCTCTGACACTGGGTCTCTGTCACCTCCATCGGAGCCAG 2160
2101 TTAACCGATCTCATAGGGTCTCTGACACTGGGTCTCTGTCACCTCCATCGGAGCCAG 2160
2161 GACCGTAGATGATTTCTTCGCCAGGACAAATCATCTGATTCAGGCTTCCTCCCA 2220
2161 GACCGTAGATGATTTCTTCGCCAGGACAAATCATCTGATTCAGGCTTCCTCCCA 2220
2221 CAAAGGCAAGACAGATTTGTTCTTAATCTGGTAGACTGTCTGGACAGAGAACTCAGA 2280
2221 CAAAGGCAAGACAGATTTGTTCTTAATCTGGTAGACTGTCTGGACAGAGAACTCAGA 2280
2281 AGTTTAAATATGCTTTGGATPAAATGATGGGTCAATGCTTCCAGAAATGGAATCC 2340
2281 AGTTTAAATATGCTTTGGATPAAATGATGGGTCAATGCTTCCAGAAATGGAATCC 2340
2341 AGTAGAATGATACCAAGATGATCTTGGTAATTTAATGCTGCTGATTCATCTCTCCCT 2400
2341 AGTAGAATGATACCAAGATGATCTTGGTAATTTAATGCTGCTGATTCATCTCTCCCT 2400
2401 GCTGAAATGAAGCACTGGCCCTGGCTTTGTTCTCCACCTTTGCTCCAATCAGAGGT 2460
2401 GCTGAAATGAAGCACTGGCCCTGGCTTTGTTCTCCACCTTTGCTCCAATCAGAGGT 2460
2461 CCATGTTTCCAGTGGATGAGAGGCGCCATCTTGAAGAGAGACCTCTCTTCCGCCCA 2520
2461 CCATGTTTCCAGTGGATGAGAGGCGCCATCTTGAAGAGAGACCTCTCTTCCGCCCA 2520
2521 CCTCTCCAGGAGCATGTTTGGAGCTTCTGAGATTTATTTCCACCAAGGAGATTTCCCA 2580
2521 CCTCTCCAGGAGCATGTTTGGAGCTTCTGAGATTTATTTCCACCAAGGAGATTTCCCA 2580
2581 GGTCCACACCTGCTCCATTTGCAATGAGAAATGCTATCCACGAGGGGTTTTCCTCCT 2640
2581 GGTCCACACCTGCTCCATTTGCAATGAGAAATGCTATCCACGAGGGGTTTTCCTCCT 2640
2641 TACCTTCCCCCAAGACCTGGATTTTTCCTCCCAACCTCCACATCTGAAGGTAGAGTGAG 2700
2641 TACCTTCCCCCAAGACCTGGATTTTTCCTCCCAACCTCCACATCTGAAGGTAGAGTGAG 2700

2701 TTCCCTCAGGTTTGATTCACCTTCAAAAGAGGCTCTACTGAACATCCAGAACCCAG 2760
2701 TTCCCTCAGGTTTGATTCACCTTCAAAAGAGGCTCTACTGAACATCCAGAACCCAG 2760
2761 CAAGAAACCTCAGCAATATTTTGTCTCTCTCAAAAGTAAATTTTGAAGTATCTCAATTTCA 2820
2761 CAAGAAACCTCAGCAATATTTTGTCTCTCTCAAAAGTAAATTTTGAAGTATCTCAATTTCA 2820
2821 GTTAAAGTAACTGCTGTTACTTAAGTATGATTAACCTTTTGTCTCAAAATGAACTTAAATGA 2880
2821 GTTAAAGTAACTGCTGTTACTTAAGTATGATTAACCTTTTGTCTCAAAATGAACTTAAATGA 2880
2881 ATTATAATCTCAGGATAGTATTTGTAATAAAGATGATTTAAATATGAATCTTATGAG 2940
2881 ATTATAATCTCAGGATAGTATTTGTAATAAAGATGATTTAAATATGAATCTTATGAG 2940
2941 TAAATATTTCAATTTTATTTAGACGGTATACTATTTCAATTTGATTAATCCATATT 3000
2941 TAAATATTTCAATTTTATTTAGACGGTATACTATTTCAATTTGATTAATCCATATT 3000
3001 ATATAAAACAAGTAGGGAGTTTATATATGTAATCTTTTCAAGTGGGAGGCTTAAATTC 3060
3001 ATATAAAACAAGTAGGGAGTTTATATATGTAATCTTTTCAAGTGGGAGGCTTAAATTC 3060
3061 TGAAGTCTGTCTTTTATGCAAGAACTGTATTTACTGTGTGTGTGTGTGTGTGTGTGTGTGT 3120
3061 TGAAGTCTGTCTTTTATGCAAGAACTGTATTTACTGTGTGTGTGTGTGTGTGTGTGTGTGT 3120
3121 TAACTTTATGCTTAAATAAATATATGTTGATTTAAATAAAAAA 3164
3121 TAACTTTATGCTTAAATAAATATATGTTGATTTAAATAAAAAA 3164

RESULT 3

US-09-744-314-14
; Sequence 14, Application US/09744314
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: BANDMAN, Olga
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: CORLEY, Neil C.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: AZIMZAI, Valda
; APPLICANT: PATTERSON, Chandra
; APPLICANT: LAL, Preeti
; APPLICANT: BAUGHN, Maziah R.
; TITLE OF INVENTION: HUMAN TROPOMYOSIN-LIKE TUMOR-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0568 PCT
; CURRENT APPLICATION NUMBER: US/09/744,314
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 09/127,665
; PRIOR FILING DATE: 1998-07-31
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq Ver. 3.0
; SEQ ID NO 14
; LENGTH: 3164
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte Clone No: 3768043
US-09-744-314-14

Query Match 100.0%; Score 3164; DB 31; Length 3164;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGTTTAAATCCATGAAAGAGCAGCAATTTTAAAGTGTATTCCACCAAAAAATAAGCTT 60
Db 1 AGGTTTAAATCCATGAAAGAGCAGCAATTTTAAAGTGTATTCCACCAAAAAATAAGCTT 60
QY 61 CAAATATGTGTATGTGAAATCTGCAGAACTAAGCGGCGGCTCAGACCAGCGCTGC 120

Db 61 CAAAATATGTGATGTGAACCTGCCGAGAACTTAAGCGGGCCGGCTCAGACACAGCGTGC 120
QY 121 CTCAGGATGAAGTGTAAACAGAGGGCCAGGGAGGTGGTGGGGACAAACATGGGCGTG 180
Db 121 CTCAGGATGAAGTGTAAACAGAGGGCCAGGGAGGTGGTGGGGACAAACATGGGCGTG 180
QY 181 TGAGGCGTGGGTCGCGCGTTCCTCCAGCTCCCGCCGAGCGCTCCAGGTCGCGTCCG 240
Db 181 TGAGGCGTGGGTCGCGCGTTCCTCCAGCTCCCGCCGAGCGCTCCAGGTCGCGTCCG 240
QY 241 CTCGGTGGTGGTCACTGCGGCAATCGGGTTCAGACCCCAAGGTCGCGTTCCTCCACC 300
Db 241 CTCGGTGGTGGTCACTGCGGCAATCGGGTTCAGACCCCAAGGTCGCGTTCCTCCACC 300
QY 301 GCTTGTGGGCGAGTGTACTGCGGTGACCGCCAGAGCGCTCCAGCGCTATGGAGGAG 360
Db 301 GCTTGTGGGCGAGTGTACTGCGGTGACCGCCAGAGCGCTCCAGCGCTATGGAGGAG 360
QY 361 CTTGGTGTACCCCTCAGCCCTACCTGGGGTGGTCTGGAGGAGCTACGAGAGTTGTG 420
Db 361 CTTGGTGTACCCCTCAGCCCTACCTGGGGTGGTCTGGAGGAGCTACGAGAGTTGTG 420
QY 421 GCAGCACTACCTGAGATGAGACAGATGAGATGATCTTATGTTTTCATCGGACTG 480
Db 421 GCAGCACTACCTGAGATGAGACAGATGAGATGATCTTATGTTTTCATCGGACTG 480
QY 481 GTGGTATGCGAGCTGTTATTCGATTTTTTGTGTTCTCTCTTTTGTGGAGAGTTTT 540
Db 481 GTGGTATGCGAGCTGTTATTCGATTTTTTGTGTTCTCTCTTTTGTGGAGAGTTTT 540
QY 541 AGATGGTATGAGTGGCTTTACGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 541 AGATGGTATGAGTGGCTTTACGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 601 GGACTAAATGAAGAT 660
Db 601 GGACTAAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 660
QY 661 GAAGGCTATGAGTGGTATCTTTAGAGGATGCGAGCTTTGAGAGAGAGAGAGAGAG 720
Db 661 GAAGGCTATGAGTGGTATCTTTAGAGGATGCGAGCTTTGAGAGAGAGAGAGAGAG 720
QY 721 GAAGCAGAGTGTGGAGCAGAGCTGTGAGAGAGTGAACAGGTGCAATCTGAACTTGG 780
Db 721 GAAGCAGAGTGTGGAGCAGAGCTGTGAGAGAGTGAACAGGTGCAATCTGAACTTGG 780
QY 781 GATGAATCTCTGTCTAGAGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 781 GATGAATCTCTGTCTAGAGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 841 GATGAATCTAGTGGGATATTTCAAGAGTATACAGTCTCTAGAGAGAGAGTGAAGATCC 900
Db 841 GATGAATCTAGTGGGATATTTCAAGAGTATACAGTCTCTAGAGAGAGAGTGAAGATCC 900
QY 901 CTCAGATCAAAATAGCTGAAGCCAAATCATCTGAGAGAGAGAGAGAGAGAGAGAGAG 960
Db 901 CTCAGATCAAAATAGCTGAAGCCAAATCATCTGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 961 CGAGGGCTATAGCATTAAGAGTGTGTTGAATGAAGAGAGAGAGAGAGAGAGAGAGAG 1020
Db 961 CGAGGGCTATAGCATTAAGAGTGTGTTGAATGAAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGGAAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db 1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGGAAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1081 CAGAAATTAACATTTGAAGAGCTCCAGAGTACAGGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 1081 CAGAAATTAACATTTGAAGAGCTCCAGAGTACAGGAGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 1141 AATCAATCAAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200

Db 1141 AATCAATCAAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 1201 GAAGAGAGACACAGGATGATTAACCTGGATTTAGAAAGTGAACAGAGTGAATCGGAAAT 1260
Db 1201 GAAGAGAGACACAGGATGATTAACCTGGATTTAGAAAGTGAACAGAGTGAATCGGAAAT 1260
QY 1261 GGTGCTTACTTATAGATTAATCTCCAAAGAGAGCTTTGAAGAGAGAGAGAGAGAGAGAG 1320
Db 1261 GGTGCTTACTTATAGATTAATCTCCAAAGAGAGCTTTGAAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 TTAATGCTTTTAAAAACCTTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 TTAATGCTTTTAAAAACCTTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 1381 GAAGTTGATTAACAAAGAGAGCTTTACAGAGAGATTTAAAAATCTTACAGAGTCAACAA 1440
Db 1381 GAAGTTGATTAACAAAGAGAGCTTTACAGAGAGATTTAAAAATCTTACAGAGTCAACAA 1440
QY 1441 GCATCTTTGAGTCAAG 1500
Db 1441 GCATCTTTGAGTCAAG 1500
QY 1501 CTTAAAGTATGATCAATTTATATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
Db 1501 CTTAAAGTATGATCAATTTATATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
QY 1561 GTAG 1620
Db 1561 GTAG 1620
QY 1621 AGCAGTCCAGTGAAG 1680
Db 1621 AGCAGTCCAGTGAAG 1680
QY 1681 TTGAG 1740
Db 1681 TTGAG 1740
QY 1741 GATTAATGTTGGTGGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
Db 1741 GATTAATGTTGGTGGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
QY 1801 GCTCAACAG 1860
Db 1801 GCTCAACAG 1860
QY 1861 CTTATGAGTCAATGTTTCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
Db 1861 CTTATGAGTCAATGTTTCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
QY 1921 TCACCATTTGGTGGGCTTCTCATCTGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
Db 1921 TCACCATTTGGTGGGCTTCTCATCTGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980
QY 1981 GAGGTCCTCACTGAG 2040
Db 1981 GAGGTCCTCACTGAG 2040
QY 2041 CCAGGAG 2100
Db 2041 CCAGGAG 2100
QY 2101 TTAACCGAG 2160
Db 2101 TTAACCGAG 2160
QY 2161 GACCGTAGGAG 2220
Db 2161 GACCGTAGGAG 2220
QY 2221 CAAAGGCAAG 2280
Db 2221 CAAAGGCAAG 2280

Db 841 GATGATTTGATGGCGGATATTTCAAAAAGTATACAGTCTCTAGAGATGAGTCAAAATCC 900
QY 901 CTCAAATCACAATAGCTGAAGCCAAAATCATCTCCAAAGACATTTAAATCAGTGAAGAA 960
Db 901 CTCAAATCACAATAGCTGAAGCCAAAATCATCTCCAAAGACATTTAAATCAGTGAAGAA 960
QY 961 CGACGGGCTATAGCAATAAAGATGCTTTGAATGAAATTTCTCAACTTCAGACAGCCAT 1020
Db 961 CGACGGGCTATAGCAATAAAGATGCTTTGAATGAAATTTCTCAACTTCAGACAGCCAT 1020
QY 1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGGAAGAGAGAGTGAAGTGAAGTGAAGTGAAGT 1080
Db 1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGGAAGAGAGAGTGAAGTGAAGTGAAGTGAAGT 1080
QY 1081 CAGAAAATAACATTTGAGACTCCAAAGTACACGACAGAACAGTTCTCAATGATTAAGAA 1140
Db 1081 CAGAAAATAACATTTGAGACTCCAAAGTACACGACAGAACAGTTCTCAATGATTAAGAA 1140
QY 1141 AATCATCATCAAGACCTGACTGGACACTTGCCTCAATGATGAAGATCAGGCTGTGTGCTT 1200
Db 1141 AATCATCATCAAGACCTGACTGGACACTTGCCTCAATGATGAAGATCAGGCTGTGTGCTT 1200
QY 1201 GAAGAGACACAAACGATGATGATACCTGGAAATGAGAGTCAACAGTGAATCGGAAT 1260
Db 1201 GAAGAGACACAAACGATGATGATACCTGGAAATGAGAGTGAATCGGAAT 1260
QY 1261 GGTGCTTACTTAGATAATCTCCAAAGAGAGCTTTGAAGAACTGATTCATGCTGCTAAG 1320
Db 1261 GGTGCTTACTTAGATAATCTCCAAAGAGAGCTTTGAAGAACTGATTCATGCTGCTAAG 1320
QY 1321 TTAATGCTTCTTTAAACCTTTAGAGAGAGAGAGTCAAGAGTCAATTTATATTCAGTGTGCT 1380
Db 1321 TTAATGCTTCTTTAAACCTTTAGAGAGAGAGAGTCAAGAGTCAATTTATATTCAGTGTGCT 1380
QY 1381 GAAGTTGATAAAACAAAGAGAGAGCTTACAGAGCATATTTAAATACTTCAGACTCAACAA 1440
Db 1381 GAAGTTGATAAAACAAAGAGAGAGCTTACAGAGCATATTTAAATACTTCAGACTCAACAA 1440
QY 1441 GCATCTTCGAGTCAGAAAACACACATTTTGAATGAGATCAGAGCTTCACAGAA 1500
Db 1441 GCATCTTCGAGTCAGAAAACACACATTTTGAATGAGATCAGAGCTTCACAGAA 1500
QY 1501 CTTAAAGTAAATGACTGAAATTTATCAAGAAAATGAAATGAAATGAAATGAAATGAAATGAA 1560
Db 1501 CTTAAAGTAAATGACTGAAATTTATCAAGAAAATGAAATGAAATGAAATGAAATGAAATGAA 1560
QY 1561 GTAGAGAAAATATTCGGTTAGAGAAAGAGAACTTTCTTAAAGTAGATGAAAGATC 1620
Db 1561 GTAGAGAAAATATTCGGTTAGAGAAAGAGAACTTTCTTAAAGTAGATGAAAGATC 1620
QY 1621 AGCCATGCCACTGAAGAGCTGGAGACTTATAGAAAGCGGCAAGATCTTGAAGAGAA 1680
Db 1621 AGCCATGCCACTGAAGAGCTGGAGACTTATAGAAAGCGGCAAGATCTTGAAGAGAA 1680
QY 1681 TTGGAGAGAACTATTCATCTTATCAAGGGCAGATTTATTTCCCATGAGAAAAGACAT 1740
Db 1681 TTGGAGAGAACTATTCATCTTATCAAGGGCAGATTTATTTCCCATGAGAAAAGACAT 1740
QY 1741 GATAATTTGTTGGCAGCTCGGAATGCTGAAAGAGAACTCAATGATTTAAGGAAAGAAAT 1800
Db 1741 GATAATTTGTTGGCAGCTCGGAATGCTGAAAGAGAACTCAATGATTTAAGGAAAGAAAT 1800
QY 1801 GCTCAACAGACAAAATTAACAGAGCTTAAATTTGAACCTTTTGAAGAAAGAT 1860
Db 1801 GCTCAACAGACAAAATTAACAGAGCTTAAATTTGAACCTTTTGAAGAAAGAT 1860
QY 1861 CTTATGCACTCGATGCTCAATATCAGATTTGGCAGAGAGATTTCCCATATGCTGCC 1920
Db 1861 CTTATGCACTCGATGCTCAATATCAGATTTGGCAGAGAGATTTCCCATATGCTGCC 1920
QY 1921 TCACCAATGGGTTGGCTTCATCTGAACAAGAGCTTTTCTCTCTCTCCCTCAACTTTGTTG 1980

Db 1921 TCACCAATGGGTTGGCTTCATCTGAACAAGAGCTTTTCTCTCTCTCCCTCAACTTTGTTG 1980
QY 1981 GAGGTCCTCACTCAGACTCTCACTTTGCTTCCAGGGGAGGAGAGGCTTCAGAGGC 2040
Db 1981 GAGGTCCTCACTCAGACTCTCACTTTGCTTCCAGGGGAGGAGAGGCTTCAGAGGC 2040
QY 2041 CCAGGGAATCCTCTGAGCAATCAGATTTACCAATGAAAGAGGAGAAATCAAGCTGTGTAGG 2100
Db 2041 CCAGGGAATCCTCTGAGCAATCAGATTTACCAATGAAAGAGGAGAAATCAAGCTGTGTAGG 2100
QY 2101 TTAACCGATCTCATAGGCTCCCTCTGACACTGGGTCTCTGTCACCTCATGGGACAG 2160
Db 2101 TTAACCGATCTCATAGGCTCCCTCTGACACTGGGTCTCTGTCACCTCATGGGACAG 2160
QY 2161 GACCGTAGGATGATTTTCTCGCCAGGACAAATCATATCTCTGATTCAGCCCTTCTCCA 2220
Db 2161 GACCGTAGGATGATTTTCTCGCCAGGACAAATCATATCTCTGATTCAGCCCTTCTCCA 2220
QY 2221 CAAAGGCAAGACAGATTTTGTCTTAATTTCTGTAGACTCTGTGACAGCAAGCTCAGA 2280
Db 2221 CAAAGGCAAGACAGATTTTGTCTTAATTTCTGTAGACTCTGTGACAGCAAGCTCAGA 2280
QY 2281 AGTTTAAATGCTTCTTTGGATAAAATGGATGGGTCAATGCTTTCAGAAATGGAATCC 2340
Db 2281 AGTTTAAATGCTTCTTTGGATAAAATGGATGGGTCAATGCTTTCAGAAATGGAATCC 2340
QY 2341 AGTAGAAATGATACCAAGATGATCTTGGTAATTTAAATTTGGCTGATTCATCTCTCCCT 2400
Db 2341 AGTAGAAATGATACCAAGATGATCTTGGTAATTTAAATTTGGCTGATTCATCTCTCCCT 2400
QY 2401 GCTGAAATCAAGCCACTGGCCCTGCTTGTTCCTCCACCTCTTCTCAATCAGAGGT 2460
Db 2401 GCTGAAATCAAGCCACTGGCCCTGCTTGTTCCTCCACCTCTTCTCAATCAGAGGT 2460
QY 2461 CCAATGTTTCCAATGATGCAAGAGCCCAATTTTGAAGAGAGAGCTTCTTTCCCTCCCA 2520
Db 2461 CCAATGTTTCCAATGATGCAAGAGCCCAATTTTGAAGAGAGAGCTTCTTTCCCTCCCA 2520
QY 2521 CCTCTCCAGAGCCATGTTTGGAGCTTCTCGAGATTTATTTTCCCAAGAGGATTTCCCA 2580
Db 2521 CCTCTCCAGAGCCATGTTTGGAGCTTCTCGAGATTTATTTTCCCAAGAGGATTTCCCA 2580
QY 2581 GGTCCACCACTGCTGCTCAATTTGCAATGAGAAATGCTATCCACCGAGGGGTTTCTCTCT 2640
Db 2581 GGTCCACCACTGCTGCTCAATTTGCAATGAGAAATGCTATCCACCGAGGGGTTTCTCTCTCT 2640
QY 2641 TACCTTCCCAAGAGCTGGATTTTCCCTCCACCTTCAATGAGCTGCTACTGAAATCAGAAC 2700
Db 2641 TACCTTCCCAAGAGCTGGATTTTCCCTCCACCTTCAATGAGCTGCTACTGAAATCAGAAC 2700
QY 2701 TTCCCTCAGGTTTGAATTCACCTTCAATGAGCTGCTACTGAAATCAGAAC 2760
Db 2701 TTCCCTCAGGTTTGAATTCACCTTCAATGAGCTGCTACTGAAATCAGAAC 2760
QY 2761 CAAAGAACCTGACAATATTTTGTCTTCAAAAGTAAATTTGATCTGATCTCATTTTCA 2820
Db 2761 CAAAGAACCTGACAATATTTTGTCTTCAAAAGTAAATTTGATCTGATCTCATTTTCA 2820
QY 2821 GTTTAAAGTAACTGCTGCTTAAAGTAACTGCTTAAAGTAACTGCTTAAAGTAACTGCTTAA 2880
Db 2821 GTTTAAAGTAACTGCTGCTTAAAGTAACTGCTTAAAGTAACTGCTTAAAGTAACTGCTTAA 2880
QY 2881 ATTATAATTTCTCAGGATAGTATTTTGTAAATGAAGTATTTAAATGAATCTTATGAG 2940
Db 2881 ATTATAATTTCTCAGGATAGTATTTTGTAAATGAAGTATTTAAATGAATCTTATGAG 2940
QY 2941 TAAATTAATTTCAATTTTATTTAGAGCTGATACTATTTCAATTTGATTAATCCACTATT 3000
Db 2941 TAAATTAATTTCAATTTTATTTAGAGCTGATACTATTTCAATTTGATTAATCCACTATT 3000
QY 3001 ATATAAACAATAGTGGAGTTTATATATGATTTCTTCAAGTGGGAGGCTTTAAATTC 3060
Db 3001 ATATAAACAATAGTGGAGTTTATATATGATTTCTTCAAGTGGGAGGCTTTAAATTC 3060

QY 3061 TGAAGTCTGCTTATCCAGAACTGATTTACTGTTGTTGGCAAAATGTGAAG 3120
Db 3061 TGAAGTCTGCTTATCCAGAACTGATTTACTGTTGTTGGCAAAATGTGAAG 3120
QY 3121 TAACTTTATGCTTAAATAAATATTAGTTGATTTAAAAA 3164
Db 3121 TAACTTTATGCTTAAATAAATATTAGTTGATTTAAAAA 3164

RESULT 5

US-60-240-409-320
; Sequence 320, Application US/60240409
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 P
; CURRENT APPLICATION NUMBER: US/60/240,409
; CURRENT FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 320
; LENGTH: 3164
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 3768043CB1
US-60-240-409-320

Query Match 100.0%; Score 3164; DB 81; Length 3164;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGTTTAAATCCATGAAGAAGACAGCAATTTTAAAGTGTATTCACCAAAAAATAAAGCTT 60
Db 1 AGGTTTAAATCCATGAAGAAGACAGCAATTTTAAAGTGTATTCACCAAAAAATAAAGCTT 60
QY 61 CAAAAATATGTATGTAAGTAACTGCGAAGTAAAGCGGGCGGGCTCAGACAGCGCTGC 120
Db 61 CAAAAATATGTATGTAAGTAACTGCGAAGTAAAGCGGGCGGGCTCAGACAGCGCTGC 120
QY 121 CTAGGATGTAAAGTGTAAACAAGAGGCCAGGGAGGTGTGGGGGACAACTGGGGCTG 180
Db 121 CTAGGATGTAAAGTGTAAACAAGAGGCCAGGGAGGTGTGGGGGACAACTGGGGCTG 180
QY 181 TGAGGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 240
Db 181 TGAGGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 240
QY 241 CTCGGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 300
Db 241 CTCGGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 300
QY 301 GCTTGTGTGGCCAGTGTACTGGGTGACCGCCAGAGAGCGCTCGACGTATGGAGGAG 360
Db 301 GCTTGTGTGGCCAGTGTACTGGGTGACCGCCAGAGAGCGCTCGACGTATGGAGGAG 360
QY 361 CCTGGTCTACCCCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 420
Db 361 CCTGGTCTACCCCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 420
QY 421 GCAGCACTACCTGAGATGATGAGACAGATGAGAAATCCTTATGTTTTCATCGGAAC 480
Db 421 GCAGCACTACCTGAGATGATGAGACAGATGAGAAATCCTTATGTTTTCATCGGAAC 480
QY 481 GTGTATGTGACGTGTATTTGATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 540
Db 481 GTGTATGTGACGTGTATTTGATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 540
QY 541 AGATCGGTTAGAGTGGGCTTACGTGGGAAGAGAGCAAAATCTTGGTGCAACGCTTCT 600
Db 541 AGATCGGTTAGAGTGGGCTTACGTGGGAAGAGAGCAAAATCTTGGTGCAACGCTTCT 600

Db 541 AGATCGGTTAGAGTGGGCTTACGTGGGAAGAGAGCAAAATCTTGGTGCAACGCTTCT 600
QY 601 GGCATTAATTCAGAAAAAATCTAACTACTTGAATAATTTAGCCCTTATTTCAAAAAGAGTAT 660
Db 601 GGCATTAATTCAGAAAAAATCTAACTACTTGAATAATTTAGCCCTTATTTCAAAAAGAGTAT 660
QY 661 GAAGGCTATGAAGTGAAGTCACTCTTTAGAGATGCCAGCTTTGAGAGAGGAGGAGCAGAA 720
Db 661 GAAGGCTATGAAGTGAAGTCACTCTTTAGAGATGCCAGCTTTGAGAGAGGAGGAGCAGAA 720
QY 721 GAAGCAGAAATTTGGAGGCAACCTGTGAAAGCTGACAGGTCCTCAATTTCTGAACCTTGG 780
Db 721 GAAGCAGAAATTTGGAGGCAACCTGTGAAAGCTGACAGGTCCTCAATTTCTGAACCTTGG 780
QY 781 GATGAAATCTCTCTCTAGAAAAAGACTTAAAAAAGAGAAATCTAAACATTTCTCAACAA 840
Db 781 GATGAAATCTCTCTCTAGAAAAAGACTTAAAAAAGAGAAATCTAAACATTTCTCAACAA 840
QY 841 GATGAAATTTGATGGCGGATATTTCAAAAAGTATACAGTCTCTAGAGATGAGTCAAAATCC 900
Db 841 GATGAAATTTGATGGCGGATATTTCAAAAAGTATACAGTCTCTAGAGATGAGTCAAAATCC 900
QY 901 CTCAAAATCACAATAGCTGAAGCCCAAAATCATCTGCAAGACATTTAAAAATGAGTGAAGAA 960
Db 901 CTCAAAATCACAATAGCTGAAGCCCAAAATCATCTGCAAGACATTTAAAAATGAGTGAAGAA 960
QY 961 CGACGGCTATAGCAATAAAGATGCTTTGATGAAATTTCTCAACTTCAGACAGCCAT 1020
Db 961 CGACGGCTATAGCAATAAAGATGCTTTGATGAAATTTCTCAACTTCAGACAGCCAT 1020
QY 1021 AAAACAGCTTTTTCAGCAAGAGCTGAAGTATGAAAGAGAGAGTGAAGTAAATAA 1080
Db 1021 AAAACAGCTTTTTCAGCAAGAGCTGAAGTATGAAAGAGAGAGTGAAGTAAATAA 1080
QY 1081 CAGAAAAATACATTTGAAGTCCAAAGTACACGAGAGAGTGAAGTGAAGTGAAGAA 1140
Db 1081 CAGAAAAATACATTTGAAGTCCAAAGTACACGAGAGAGTGAAGTGAAGTGAAGAA 1140
QY 1141 AATCACTCAAGACCTGACTGGACATTTGCCAATGATGAAGATCAGGCTGCTGCTG 1200
Db 1141 AATCACTCAAGACCTGACTGGACATTTGCCAATGATGAAGATCAGGCTGCTGCTG 1200
QY 1201 GAAGAGACACAAACGATGATGAATTAACCTGGAATTTAGAAAGTGAACAGTGAATCGGAAAT 1260
Db 1201 GAAGAGACACAAACGATGATGAATTAACCTGGAATTTAGAAAGTGAACAGTGAATCGGAAAT 1260
QY 1261 GGTGCTTACTAGATAATCTTCAAAAAGAGCTTTGAAGAACTGATTCATGCTGCTAAG 1320
Db 1261 GGTGCTTACTAGATAATCTTCAAAAAGAGCTTTGAAGAACTGATTCATGCTGCTAAG 1320
QY 1321 TTAATGCTTCTTTAAAAACCTTAGAAGGAGAAAGAAACCAAAATTTATATTCAGTTGTCT 1380
Db 1321 TTAATGCTTCTTTAAAAACCTTAGAAGGAGAAAGAAACCAAAATTTATATTCAGTTGTCT 1380
QY 1381 GAAGTTGATAAACAAGAGAGCTTACAGAGCATTTAAAAATCTTCAGACTCAACAA 1440
Db 1381 GAAGTTGATAAACAAGAGAGCTTACAGAGCATTTAAAAATCTTCAGACTCAACAA 1440
QY 1441 GCATCTTTGCACTCAGAAAAACACATTTTGAATAAGAGATCAGAACTTCAACAGAAA 1500
Db 1441 GCATCTTTGCACTCAGAAAAACACATTTTGAATAAGAGATCAGAACTTCAACAGAAA 1500
QY 1501 CTTAAGTAATGCTGAATTTATCAAGAAATGAAATGAAATCTCCAGGAAATTAACA 1560
Db 1501 CTTAAGTAATGCTGAATTTATCAAGAAATGAAATGAAATCTCCAGGAAATTAACA 1560
QY 1561 GTAGAGGAAAAATTTATCGGTTAGAGAAAGAGAAATCTTAAAAAGTGAAGAAATGATC 1620
Db 1561 GTAGAGGAAAAATTTATCGGTTAGAGAAAGAGAAATCTTAAAAAGTGAAGAAATGATC 1620
QY 1621 AGCCATGCCACTGAAGAGCTGGAGACCTATAGAAAGAGGAGCAGAAAGATCTTGAAGAGAA 1680
Db 1621 AGCCATGCCACTGAAGAGCTGGAGACCTATAGAAAGAGGAGCAGAAAGATCTTGAAGAGAA 1680

1681 TTGGAGAACTATTCTATCTATCAAGGCGAGATTATTTCCCATGAGAAAGACACAT 1740
1681 TTGGAGAACTATTCTATCTATCAAGGCGAGATTATTTCCCATGAGAAAGACACAT 1740
1741 GATAATTGGTTGGCAGCTCGGAATCCTGAAAGAAACCTCAATGATTTAAGGAAAGAAAT 1800
1741 GATAATTGGTTGGCAGCTCGGAATCCTGAAAGAAACCTCAATGATTTAAGGAAAGAAAT 1800
1801 GCTCAACAGACAAATAATTAACTCAACAGAGCTTAAATTTGAACCTTTTGAAGAAAGAT 1860
1801 GCTCAACAGACAAATAATTAACTCAACAGAGCTTAAATTTGAACCTTTTGAAGAAAGAT 1860
1861 CTTATGCACTCGATGTTCCAAATACAGCAATTTGCACAGAGCAATCCCATATAGTCC 1920
1861 CTTATGCACTCGATGTTCCAAATACAGCAATTTGCACAGAGCAATCCCATATAGTCC 1920
1921 TCACCAATTTGGTGGCTTCATCTGAACAGAGCTTTCTCTCTCCCACTTTGTTG 1980
1921 TCACCAATTTGGTGGCTTCATCTGAACAGAGCTTTCTCTCTCCCACTTTGTTG 1980
1981 GAGGTCCTCACTCAGACTCTCACCTTTGTTCCAGGGGAGGAGGAGGCTCACGAGC 2040
1981 GAGGTCCTCACTCAGACTCTCACCTTTGTTCCAGGGGAGGAGGAGGCTCACGAGC 2040
2041 CCAGGAATCCTCTGACCATCAGATTCAATGAAGAGGAGGAGGAGGCTCACGAGC 2100
2041 CCAGGAATCCTCTGACCATCAGATTCAATGAAGAGGAGGAGGAGGCTCACGAGC 2100
2101 TTAACCGATCCTCATAGGCTCCCTCTGACACTGGGCTCTCTCTCACTCCATGGACACG 2160
2101 TTAACCGATCCTCATAGGCTCCCTCTGACACTGGGCTCTCTCTCACTCCATGGACACG 2160
2161 GACCGTAGATGATGTTCTCGCCAGACCAATCATATCCTGATTCAGCCCTTCTCTCA 2220
2161 GACCGTAGATGATGTTCTCGCCAGACCAATCATATCCTGATTCAGCCCTTCTCTCA 2220
2221 CAAAGGCAACAGAGATTTGTTCTAATCTGTTAGACTGTTGGACCAAGCAATCTCAG 2280
2221 CAAAGGCAACAGAGATTTGTTCTAATCTGTTAGACTGTTGGACCAAGCAATCTCAG 2280
2281 AGTTTAAATAGCTCTTTGGATTAATAGGATGGGTCATGCTTCAGAAATGGATCC 2340
2281 AGTTTAAATAGCTCTTTGGATTAATAGGATGGGTCATGCTTCAGAAATGGATCC 2340
2341 AGTAGAAATCATACCAAGATGATCTTGGTAAATTTAAATGCTGATTCATCTCTCCCT 2400
2341 AGTAGAAATCATACCAAGATGATCTTGGTAAATTTAAATGCTGATTCATCTCTCCCT 2400
2401 GCTGAAATCAAGCCACTGCGCTGCTTTGTTCTCTCCCACTTGTCTCAATCAGAGT 2460
2401 GCTGAAATCAAGCCACTGCGCTGCTTTGTTCTCTCCCACTTGTCTCAATCAGAGT 2460
2461 CCATGTTTCCAGTGGATCAAGAGGCCCAATCTTGAAGAGAGGACTCTCTTCCCCCA 2520
2461 CCATGTTTCCAGTGGATCAAGAGGCCCAATCTTGAAGAGAGGACTCTCTTCCCCCA 2520
2521 CCTCCTCCAGAGCCACTGTTGGAGCTTCTCGAGATTAATTTTCCACCAAGGATTTCCCA 2580
2521 CCTCCTCCAGAGCCACTGTTGGAGCTTCTCGAGATTAATTTTCCACCAAGGATTTCCCA 2580
2581 GGTCCACCACTGCTCCATTTGCAATCAGAAATGCTCTATCCAGGAGGCTTTCTCTCT 2640
2581 GGTCCACCACTGCTCCATTTGCAATCAGAAATGCTCTATCCAGGAGGCTTTCTCTCT 2640
2641 TACCTTCCCCCAAGACCTGGATTTTCCCCCACTTCTGAGGTAGAGTGGAG 2700
2641 TACCTTCCCCCAAGACCTGGATTTTCCCCCACTTCTGAGGTAGAGTGGAG 2700
2701 TTCCCTCAGGTTGATTCACCTTCAATCAGAGCTCTCTGAACTCAGACACAG 2760
2701 TTCCCTCAGGTTGATTCACCTTCAATCAGAGCTCTCTGAACTCAGACACAG 2760

2761 CAAGAAACCTGCAATATTTTGTCTCTCTCAAAAGTAATTTTGTGACTGATCTCATTTCA 2820
2761 CAAGAAACCTGCAATATTTTGTCTCTCTCTCAAAAGTAATTTTGTGACTGATCTCATTTCA 2820
2821 GTTTAAGTAACCTGCTGTTAAAGTGATTAACACTTTTGTCTCAAAATTTGAAGCTTAATGA 2880
2821 GTTTAAGTAACCTGCTGTTAAAGTGATTAACACTTTTGTCTCAAAATTTGAAGCTTAATGA 2880
2881 ATTATATTTCTCAGGATAGTATTTTGTAAATTAAGATGATTTAAATATGAATCTTATGAG 2940
2881 ATTATATTTCTCAGGATAGTATTTTGTAAATTAAGATGATTTAAATATGAATCTTATGAG 2940
2941 TAAATTTATTTCAATTTTATTTAGACGGTATTAACATTTCAATTTGATTAATCCACTATT 3000
2941 TAAATTTATTTCAATTTTATTTAGACGGTATTAACATTTCAATTTGATTAATCCACTATT 3000
3001 ATATAACAATAGTGGAGTTTATATATATGTAATCTTTTCAAGTGGGAGGCTTTAAATTC 3060
3001 ATATAACAATAGTGGAGTTTATATATATGTAATCTTTTCAAGTGGGAGGCTTTAAATTC 3060
3061 TGAAGTCTGCTCTTTATGCCAAGAACTGATTTACTGTTGTTGTCACAAATGTGAAAG 3120
3061 TGAAGTCTGCTCTTTATGCCAAGAACTGATTTACTGTTGTTGTCACAAATGTGAAAG 3120
3121 TAACTTTATGCTTTAAATAAATTAATGATTTGATTTTAAATAAAAAA 3164
3121 TAACTTTATGCTTTAAATAAATTAATGATTTGATTTTAAATAAAAAA 3164

RESULT 6

US-60-324-185-1645
; Sequence 1645, Application US/60324185
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Disp, Diph
; TITLE OF INVENTION: METHOD FOR THE IDENTIFICATION OF SEQUENCE POLYMORPHISMS USING
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCE DATABASES, AND SINGLE NUCLEOTIDE
; FILE REFERENCE: GX-0019-1 P
; CURRENT APPLICATION NUMBER: US/60/324,185
; NUMBER OF SEQ ID NOS: 35862
; SOFTWARE: PERL Program
; SEQ ID NO 1645
; LENGTH: 3164
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 039946.18
US-60-324-185-1645

Query Match 99.8%; Score 3159.2; DB 89; Length 3164;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3161; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 AGGTTTAAATCCATGAGAGACGCAATTTTAAAGTGATTTACCAAAATAAAGCTT 60
1 AGGTTTAAATCCATGAGAGACGCAATTTTAAAGTGATTTACCAAAATAAAGCTT 60
61 CAAAAATATGTGATGTGAAAACTGCCAGAACTAAGCGCGGCGCGGCTCAGACGCGTGC 120
61 CAAAAATATGTGATGTGAAAACTGCCAGAACTAAGCGCGGCGCGGCTCAGACGCGTGC 120
121 CTCAGATGTAAAGTGTAAACAGAGGCGGAGGAGGTTGGGGGACAAATGGGCTG 180
121 CTCAGATGTAAAGTGTAAACAGAGGCGGAGGAGGTTGGGGGACAAATGGGCTG 180
181 TGAGGCGCTGTGGGTGCCCGGCTTCCCGAGCTCCCCCGAGCCCGCTCCACAGTGGTCCG 240
181 TGAGGCGCTGTGGGTGCCCGGCTTCCCGAGCTCCCCCGAGCCCGCTCCACAGTGGTCCG 240

QY 241 CTCGGTGTGTTGTCAGCTGGCATTCCGGTTCCAGACCCAGGCTCGGTGTTCTCCACC 300
Db 241 CTCGGTGTGTTGTCAGCTGGCATTCCGGTTCCAGACCCAGGCTCGGTGTTCTCCACC 300
QY 301 GCTTGTGTGGCCHGTGTTACTCGGTGACCGCCAGAGCAGCTCGACGCTATGGAGGAG 360
Db 301 GCTTGTGTGGCCHGTGTTACTCGGTGACCGCCAGAGCAGCTCGACGCTATGGAGGAG 360
QY 361 CTTGTGCTTACCCTCAGCCCTACCTCGGCTGGTCTCGGAGGAGTACCGAGGTTGTG 420
Db 361 CTTGTGCTTACCCTCAGCCCTACCTCGGCTGGTCTCGGAGGAGTACCGAGGTTGTG 420
QY 421 GCAGCACTACTGAGAGTATGAGACCAAGATGAGAACTCTTTATGGTTTGCATCGGAATG 480
Db 421 GCAGCACTACTGAGAGTATGAGACCAAGATGAGAACTCTTTATGGTTTGCATCGGAATG 480
QY 481 GTGTAATGTCAGCTGTATTTGTAATTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
Db 481 GTGTAATGTCAGCTGTATTTGTAATTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
QY 541 AGATCGGTTAGGAGTGGCTTTTACGTGGGAGAGAGCAAAAATCTGGTGCACCGCTTTCT 600
Db 541 AGATCGGTTAGGAGTGGCTTTTACGTGGGAGAGAGCAAAAATCTGGTGCACCGCTTTCT 600
QY 601 GGACTAATGAAGAAATCTAACTACTTGAATAATTTAGCCCTTATCAAAAAGAGTAT 660
Db 601 GGACTAATGAAGAAATCTAACTACTTGAATAATTTAGCCCTTATCAAAAAGAGTAT 660
QY 661 GAAGGCTATGAAGTAGAGTCACTTTAGAGGATGCCAGCTTTGAGAGGAGGAGCAGAGAA 720
Db 661 GAAGGCTATGAAGTAGAGTCACTTTAGAGGATGCCAGCTTTGAGAGGAGGAGCAGAGAA 720
QY 721 GAAGCAGAGTGTGGAGGCAACTGTGAAGAGCTGAACAGGTCCTAACTCTGAGCTTGAG 780
Db 721 GAAGCAGAGTGTGGAGGCAACTGTGAAGAGCTGAACAGGTCCTAACTCTGAGCTTGAG 780
QY 781 GATGAATCTCTGCTCAGAAAAGAGCTTAAACACAGAGAACTTAAACATCTCAACAA 840
Db 781 GATGAATCTCTGCTCAGAAAAGAGCTTAAACACAGAGAACTTAAACATCTCAACAA 840
QY 841 GATGAATGATGGCGATATTTCAAAAGATATACAGTCTCTAGAGATGATGATCAAAATCC 900
Db 841 GATGAATGATGGCGATATTTCAAAAGATATACAGTCTCTAGAGATGATGATCAAAATCC 900
QY 901 CTCAAATCACAATAGCTGAAGCAGCAATCATCTGCAAGACATTTAAATGAGTGAAGAA 960
Db 901 CTCAAATCACAATAGCTGAAGCAGCAATCATCTGCAAGACATTTAAATGAGTGAAGAA 960
QY 961 CGAGCGCTATAGCAATAAAGATGCTTTGAATGAATAATTTCTCAATTCAGACAAGCCAT 1020
Db 961 CGAGCGCTATAGCAATAAAGATGCTTTGAATGAATAATTTCTCAATTCAGACAAGCCAT 1020
QY 1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGGAAGGAGAGTGAAGTGAATTAATAA 1080
Db 1021 AAACAGCTTTTTCAGCAAGAGCTGAAGTATGGAAGGAGAGTGAAGTGAATTAATAA 1080
QY 1081 CAGAAAATAACATTTGAAGACTCCAAAGTACACGAGAACAGTTCTGAATGAATAAAGAA 1140
Db 1081 CAGAAAATAACATTTGAAGACTCCAAAGTACACGAGAACAGTTCTGAATGAATAAAGAA 1140
QY 1141 AATCAGATCAGACCTGACCTGACACTTCCCAATGATGAAGATCAGGCTGCTGCTT 1200
Db 1141 AATCAGATCAGACCTGACCTGACACTTCCCAATGATGAAGATCAGGCTGCTGCTT 1200
QY 1201 GAAGAAGACACACCGGATGATGATAACCTGGAATTAGAAGTGAACAGTGAATCGGAAT 1260
Db 1201 GAAGAAGACACACCGGATGATGATAACCTGGAATTAGAAGTGAACAGTGAATCGGAAT 1260
QY 1261 GGTGCTTACTTAGATAATCTCCAAAGAGCTTTGAAGAACTGATTCATGCTGCTAAG 1320
Db 1261 GGTGCTTACTTAGATAATCTCCAAAGAGCTTTGAAGAACTGATTCATGCTGCTAAG 1320
QY 1321 TTAAATGCTTCTTTAAACACCTTAGAGGAGAGAAACCAAAATTTATATTTCAGTTGCT 1380

Db 1321 TTAAATGCTTCTTTAAACACCTTAGAGGAGAGAAACCAAAATTTATATTTCAGTTGCT 1380
QY 1381 GAAGTTGATAAACAAGAGAGCTTACAGAGCATATTAATAATCTTCAGACTCAACAA 1440
Db 1381 GAAGTTGATAAACAAGAGAGCTTACAGAGCATATTAATAATCTTCAGACTCAACAA 1440
QY 1441 GCATCTTTGCACTCAGAAAACACACATTTTGAATAAGATCAGAACTTCAACAGAAA 1500
Db 1441 GCATCTTTGCACTCAGAAAACACACATTTTGAATAAGATCAGAACTTCAACAGAAA 1500
QY 1501 CTTAAAGTAATGACTGAATTAATCAAGAAATGAATGAATCTCCAGAGAAATTAACA 1560
Db 1501 CTTAAAGTAATGACTGAATTAATCAAGAAATGAATGAATCTCCAGAGAAATTAACA 1560
QY 1561 GTAGAGGAAAATTTATCGGTTAGAGAAAGAGAAAATCTTTCTAAAGTAGATGAAGATC 1620
Db 1561 GTAGAGGAAAATTTATCGGTTAGAGAAAGAGAAAATCTTTCTAAAGTAGATGAAGATC 1620
QY 1621 AGCCATGCTGAGAGCTGAGACCTTATAGAAAGCGAGCCAAAGATCTTTGAAGAGAA 1680
Db 1621 AGCCATGCTGAGAGCTGAGACCTTATAGAAAGCGAGCCAAAGATCTTTGAAGAGAA 1680
QY 1681 TTGAGAGAACTTATTCATTTATCAAGGCGAGATTTTCCCATGAGAAAAGACAT 1740
Db 1681 TTGAGAGAACTTATTCATTTATCAAGGCGAGATTTTCCCATGAGAAAAGACAT 1740
QY 1741 GATAATTCGTTGGAGCTCGGAATGCTGAAAGAACTCTCAATGATTTAAGAAAGAAAAT 1800
Db 1741 GATAATTCGTTGGAGCTCGGAATGCTGAAAGAACTCTCAATGATTTAAGAAAGAAAAT 1800
QY 1801 GCTCACAACAGACAAAATTAACCTGAAACAGAGCTTTAAATTTGAACCTTTAGAAAAGAT 1860
Db 1801 GCTCACAACAGACAAAATTAACCTGAAACAGAGCTTTAAATTTGAACCTTTAGAAAAGAT 1860
QY 1861 CTTATGCACTCGATGTTTCCAAATACAGATTTGGCAGAGAGCATTTCCCATATGCTGCC 1920
Db 1861 CTTATGCACTCGATGTTTCCAAATACAGATTTGGCAGAGAGCATTTCCCATATGCTGCC 1920
QY 1921 TCACCATTTGGGTTGGCTTCATCTGAAACAGAGCTTTCTCTCTCTCTCTCTCTCTCT 1980
Db 1921 TCACCATTTGGGTTGGCTTCATCTGAAACAGAGCTTTCTCTCTCTCTCTCTCTCTCT 1980
QY 1981 GAGGCTCCTCAGACTCTCAGCTTTGCTTCCAGGGGAGGAGGAGGCTCAGAGGC 2040
Db 1981 GAGGCTCCTCAGACTCTCAGCTTTGCTTCCAGGGGAGGAGGAGGCTCAGAGGC 2040
QY 2041 CCAGGGAATCCTTCGACCATCAGATTTACCAATGAAGAGGAGATCAAGCTGTGATAGG 2100
Db 2041 CCAGGGAATCCTTCGACCATCAGATTTACCAATGAAGAGGAGATCAAGCTGTGATAGG 2100
QY 2101 TTAAACCGATCCTCAGAGGCTCCCTCTGACACTGGGTCTCTGTCACCTCCATGGAGCCAG 2160
Db 2101 TTAAACCGATCCTCAGAGGCTCCCTCTGACACTGGGTCTCTGTCACCTCCATGGAGCCAG 2160
QY 2161 GACCGTAGATGATGTTTCTTCGCGCAGAGCAATCATATCTGATTCAGCCCTCTCTCA 2220
Db 2161 GACCGTAGATGATGTTTCTTCGCGCAGAGCAATCATATCTGATTCAGCCCTCTCTCA 2220
QY 2221 CAAAGGCAAGACAGATTTGTTCTTAATTTCTGTAGACTGTCTGGACCAAGAGAACTCAGA 2280
Db 2221 CAAAGGCAAGACAGATTTGTTCTTAATTTCTGTAGACTGTCTGGACCAAGAGAACTCAGA 2280
QY 2281 AGTTTAAATATGCTTCTTTGGATTAATGAATGGGTCAATGCTTTCAGAAATGAATTC 2340
Db 2281 AGTTTAAATATGCTTCTTTGGATTAATGAATGGGTCAATGCTTTCAGAAATGAATTC 2340
QY 2341 AGTAGAATGATACCAAGAGATCTTGGTAAATTTAAATGTCCTGATTCATCTCTCCT 2400
Db 2341 AGTAGAATGATACCAAGAGATCTTGGTAAATTTAAATGTCCTGATTCATCTCTCCT 2400
QY 2401 GCTGAAAATGAAGCCACTGGCCCTGGCTTTGTTCTCTCCACTTTGCTTCCCAATCAGAGT 2460

2401 GGTGAAATGAAGCCACTGGCCCTGGCTTTGTTCTCCACCTCTTGTCTCCAAATCAGAGGT 2460
2461 CCATTTCTTCCAGTGGATCAAGAGAGCCCATTTCTTGAGAAGAGGACCTCTTCCCCCCCA 2520
2461 CCATTTCTTCCAGTGGATCAAGAGAGCCCATTTCTTGAGAAGAGGACCTCTTCCCCCCCA 2520
2521 CCTCTCCAGGAGCCATGTTGGAGCTTCTCAGATTAATTTCCACCAAGGATTTCCCA 2580
2521 CCTCTCCAGGAGCCATGTTGGAGCTTCTCAGATTAATTTCCACCAAGGATTTCCCA 2580
2581 GGTCCACCACTGCTCCATTTGCAATGAGAAATGCTATCCACGAGGGGTTTCTCTCT 2640
2581 GGTCCACCACTGCTCCATTTGCAATGAGAAATGCTATCCACGAGGGGTTTCTCTCT 2640
2641 TACCTTCCCCCAAGACCTGGATTTTCCCCCAAGGAGGATTTCTGAGAGTGAAGTGA 2700
2641 TACCTTCCCCCAAGACCTGGATTTTCCCCCAAGGAGGATTTCTGAGAGTGAAGTGA 2700
2701 TTTCCCTCAGGTTTGAATTCACCTTCAATGAGCCTGCTACTGACATCCAGAACACAG 2760
2701 TTTCCCTCAGGTTTGAATTCACCTTCAATGAGCCTGCTACTGACATCCAGAACACAG 2760
2761 CAAGAACTGACAAATATTTTGTCTCTCTTCAAAAGTAAATTTGACTGATCTCATTTTCA 2820
2761 CAAGAACTGACAAATATTTTGTCTCTCTTCAAAAGTAAATTTGACTGATCTCATTTTCA 2820
2821 GTTTAAGTAACTGCTGTTACTTAAAGTGAATACATTTTCTCAAAATGGAAGCTTAATGGA 2880
2821 GTTTAAGTAACTGCTGTTACTTAAAGTGAATACATTTTCTCAAAATGGAAGCTTAATGGA 2880
2881 ATTATAATCTCAGGATAGTATTTTCTTAAATGAGTAAATGATGATGATGATGATGATGAT 2940
2881 ATTATAATCTCAGGATAGTATTTTCTTAAATGAGTAAATGATGATGATGATGATGATGAT 2940
2941 TAAATTTATTTCAATTTTATTTAGACGATTAATTTCAATTTGATTTGATTTGATTTGATTT 3000
2941 TAAATTTATTTCAATTTTATTTAGACGATTAATTTCAATTTGATTTGATTTGATTTGATTT 3000
3001 ATATAACAAATAGTGGAGTTTATATATGATTAATCTTTCCAGTGGGAGGCTTTAAATTC 3060
3001 ATATAACAAATAGTGGAGTTTATATATGATTAATCTTTCCAGTGGGAGGCTTTAAATTC 3060
3061 TGAAGTCTGTGCTTTATCCAGAACTGATTTACTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3120
3061 TGAAGTCTGTGCTTTATCCAGAACTGATTTACTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3120
3121 TAACTTTATGCTTAAATAAATTTATAGTTGATTTTAAATTTTTTTTTTTTTTTTTTTTTTT 3164
3121 TAACTTTATGCTTAAATAAATTTATAGTTGATTTTAAATTTTTTTTTTTTTTTTTTTTTTT 3164

RESULT 7
US-60-172-360-25590
; Sequence 25590, Application US/60172360
; GENERAL INFORMATION:
; APPLICANT: Moris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; FILE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide Polymorp
; FILE REFERENCE: GX-0007 P
; CURRENT APPLICATION NUMBER: US/60/172,360
; CURRENT FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 29838
; SOFTWARE: PERL Program
; SEQ ID NO 25590
; LENGTH: 3165
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 039946.4
US-60-172-360-25590

Query Match 99.7%; Score 3153; DB 73; Length 3165;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3164; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy 1 AGTTTAAATCCATGAAGAGACAGCAATTTTAAAGTGTATTCCACCAAAATAAAGCTT 60
Db 1 AGTTTAAATCCATGAAGAGACAGCAATTTTAAAGTGTATTCCACCAAAATAAAGCTT 60
Qy 61 CAAATAATATGTATGTGAAATCTCCAGAACTAAGCGGGCGGGCTCAGACCAAGCCTGC 120
Db 61 CAAATAATGTATGTGAAATCTCCAGAACTAAGCGGGCGGGCTCAGACCAAGCCTGC 120
Qy 121 CTGAGGATGTAAAGTGTAAAGAGGCGGAGGAGTGTGGGGGACACATGGGCGCTG 180
Db 121 CTGAGGATGTAAAGTGTAAAGAGGCGGAGGAGTGTGGGGGACACATGGGCGCTG 180
Qy 181 TGAGGCGCTGTGGGTGCGCGCTTCCCGAGCTCCCGCGAGCCCGCTCCACAGTGTGCG 240
Db 181 TGAGGCGCTGTGGGTGCGCGCTTCCCGAGCTCCCGCGAGCCCGCTCCACAGTGTGCG 240
Qy 241 CTCGGTGTGTGTGTCAGTGCGCATTTCCGGTTCAGAGCCCAAGGCTGCGTGTCTCCAC 300
Db 241 CTCGGTGTGTGTGTCAGTGCGCATTTCCGGTTCAGAGCCCAAGGCTGCGTGTCTCCAC 300
Qy 301 GCTTGTGTGTCGAGTGTACTCGGTGACCGCCAGACAGACGCTCGACGCTATGGAGGAG 360
Db 301 GCTTGTGTGTCGAGTGTACTCGGTGACCGCCAGACAGACGCTCGACGCTATGGAGGAG 360
Qy 361 CTTGGTCTACCCCTCAGCCCTACCTGGGCTGGTCTCGAGGAGCTACGACAGTGTG 420
Db 361 CTTGGTCTACCCCTCAGCCCTACCTGGGCTGGTCTCGAGGAGCTACGACAGTGTG 420
Qy 421 GCAGCACTACTCTGAGAGTATGAGACAGATGAGAAATCTTATGTTTTCATCGGAATCTG 480
Db 421 GCAGCACTACTCTGAGAGTATGAGACAGATGAGAAATCTTATGTTTTCATCGGAATCTG 480
Qy 481 GTGATATGTGAGCTGTATTTGATTTTGTGTTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
Db 481 GTGATATGTGAGCTGTATTTGATTTTGTGTTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
Qy 541 AGATCGGTGTAGGAGTGGCTTTACGTGGGAGAGAGCAAAATCTGGTGCAACGCTTTCT 600
Db 541 AGATCGGTGTAGGAGTGGCTTTACGTGGGAGAGAGCAAAATCTGGTGCAACGCTTTCT 600
Qy 601 GGACTAAATGAAGAAATCTAACTACTTTGAAATTTAGCCTTATTTCAAAAGAGTAT 660
Db 601 GGACTAAATGAAGAAATCTAACTACTTTGAAATTTAGCCTTATTTCAAAAGAGTAT 660
Qy 661 GAAGCTATGAAGTATGAGTATCTTTAGAGATGCGCAGCTTTGAGAGGAGGAGCAGAGAA 720
Db 661 GAAGCTATGAAGTATGAGTATCTTTAGAGATGCGCAGCTTTGAGAGGAGGAGCAGAGAA 720
Qy 721 GAAGCAGAAATTTGGAGGCAACCTGTGAAAAGCTGAACAGGTCCTCAATTTCTGAATTTGAG 780
Db 721 GAAGCAGAAATTTGGAGGCAACCTGTGAAAAGCTGAACAGGTCCTCAATTTCTGAATTTGAG 780
Qy 781 GATGAAATCTCTCTCTAGAAAAGTAAATCAACAGAGAAATCTTAAACATTTCTCAACAA 840
Db 781 GATGAAATCTCTCTCTAGAAAAGTAAATCAACAGAGAAATCTTAAACATTTCTCAACAA 840
Qy 841 GATGAAATTTGATGGCGGATTTTCAAAAGTATACAGTCTCTAGAGATGAGTCAAAATTC 900
Db 841 GATGAAATTTGATGGCGGATTTTCAAAAGTATACAGTCTCTAGAGATGAGTCAAAATTC 900
Qy 901 CTCAAATCAAAATAGCTGAGCCAAATCATCTGCAAGACATTTAAATGAGTGAAGAA 960
Db 901 CTCAAATCAAAATAGCTGAGCCAAATCATCTGCAAGACATTTAAATGAGTGAAGAA 960
Qy 961 CGACGGGCTATAGCAATAAAGATGCTTTTGAATGAAATTTCTCAACTTCAGACCAAGCCAT 1020
Db 961 CGACGGGCTATAGCAATAAAGATGCTTTTGAATGAAATTTCTCAACTTCAGACCAAGCCAT 1020

1021 AAAACAGCTTTTTCAGCAAGAGCTGAAGTATGAAAGGAGAGTGAAGTAAATTAATAA 1080
1021 AAAACAGCTTTTTCAGCAAGAGCTGAAGTATGAAAGGAGAGTGAAGTAAATTAATAA 1080
1081 CAGAAATTAACATTTGAAGACTCCAAAGTACACGAGAGAACTGCTGAATGATAAGAA 1140
1081 CAGAAATTAACATTTGAAGACTCCAAAGTACACGAGAGAACTGCTGAATGATAAGAA 1140
1141 AATCAGATCAAGACCTGACTGACACTTGCACACTTGCACAAATGATGAAGATCAGGCTGTGCTT 1200
1141 AATCAGATCAAGACCTGACTGACACTTGCACAAATGATGAAGATCAGGCTGTGCTT 1200
1201 GAAGAGACACAAACGAGTATGATAAATCTGGAAATGAAAGTGAACAGTGAATCGGAAAT 1260
1201 GAAGAGACACAAACGAGTATGATAAATCTGGAAATGAAAGTGAACAGTGAATCGGAAAT 1260
1261 GGTGCTTACTTATGATTAATCTCCAAAGAGGCTTTGAAGAACTGATTAATGCTGCTAAG 1320
1261 GGTGCTTACTTATGATTAATCTCCAAAGAGGCTTTGAAGAACTGATTAATGCTGCTAAG 1320
1321 TTAATGCTTCTTTTAAACCTTTAGAGAGAGAGAAACCAAAATTTATATTCAGTTGCTT 1380
1321 TTAATGCTTCTTTTAAACCTTTAGAGAGAGAGAAACCAAAATTTATATTCAGTTGCTT 1380
1381 GAAGTTGATTAACAAAGAGAGGCTTACAGAGCATATTAATAAT-CTTCAGACTCAACA 1439
1381 GAAGTTGATTAACAAAGAGAGGCTTACAGAGCATATTAATAATCTTCAGACTCAACA 1440
1440 AGCATCTTTGCAGTCAGAGAAACACACATTTTGAANAATGAGAATCAGAAGCTTCAACAGAA 1499
1441 AGCATCTTTGCAGTCAGAGAAACACACATTTTGAANAATGAGAATCAGAAGCTTCAACAGAA 1500
1500 ACTTAAGTATGACTGAATATATCAAGAAATGAANAATGAACCTCCACAGAGAAATTAAC 1559
1501 ACTTAAGTATGACTGAATATATCAAGAAATGAANAATGAACCTCCACAGAGAAATTAAC 1560
1560 AGTAGAGAGAAATTTATCGGTTAGAGAGAGAGAGAACTTTCTAAAGTAGATGAAGAGAT 1619
1561 AGTAGAGAGAAATTTATCGGTTAGAGAGAGAGAGAACTTTCTAAAGTAGATGAAGAGAT 1620
1620 CAGCCATGCCACTGAAGAGCTGGAGACTATAGAAAGCGAGCCAAAGATCTTGAAGAGA 1679
1621 CAGCCATGCCACTGAAGAGCTGGAGACTATAGAAAGCGAGCCAAAGATCTTGAAGAGA 1680
1680 ATTGGAGAGAACTATTATCTTATCAAGGCGAGATTTCCCATCAGAGAAAGAGCACA 1739
1681 ATTGGAGAGAACTATTATCTTATCAAGGCGAGATTTCCCATCAGAGAAAGAGCACA 1740
1740 TGATAATTTGGTTGGCAGCTCGGAATGCTGAAAGAAACCTCAATGATTTAAGGAAAGAAA 1800
1800 TGCTCACAACAGACAAATAATTAAGTGAACAGAGCTTAATTTCAACTTTTAGAAGAAAGA 1859
1801 TGCTCACAACAGACAAATAATTAAGTGAACAGAGCTTAATTTCAACTTTTAGAAGAAAGA 1860
1860 TCCTTATGCACTCGATGTTCCAAATACAGCAATTTGGCAGAGAGATTTCCCATATGCTGCC 1919
1861 TCCTTATGCACTCGATGTTCCAAATACAGCAATTTGGCAGAGAGATTTCCCATATGCTGCC 1920
1920 CTCACCAATGGTGGCTTCATCTGAACAGAGCTTTTCTCTCTCTCTCTCTCTCTCTCTCT 1979
1921 CTCACCAATGGTGGCTTCATCTGAACAGAGCTTTTCTCTCTCTCTCTCTCTCTCTCTCT 1980
1980 GGAGGCTCCACTCAGACTCTCACCTTTGCTTCCAGGGGAGAGAGAGGCTCACGAGG 2039
1981 GGAGGCTCCACTCAGACTCTCACCTTTGCTTCCAGGGGAGAGAGAGGCTCACGAGG 2040
2040 CCCAGGGAATCTCTGGACCATCAGATTAACCAATGAAGAGAGAGATCAAGCTGTGATAG 2099
2041 CCCAGGGAATCTCTGGACCATCAGATTAACCAATGAAGAGAGAGATCAAGCTGTGATAG 2100
2100 GTTAACCGATCTCATAGGCTCCCTCTGACACTGGGTCTCTGTCACCTCCATCGGACCA 2159

RESULT 8

2101 GTTAACCGATCTCATAGGCTCCCTCTGACACTGGGTCTCTGTCACTCCATCGGACCA 2160
2160 GGACGCTAGAGTATGATTTCTCTCGCCAGGACCAATCATATCTCTGATTGAGCCCTTCCTCC 2219
2161 GGACGCTAGAGTATGATTTCTCTCGCCAGGACCAATCATATCTCTGATTGAGCCCTTCCTCC 2220
2220 ACAAAGCACAAGACAGATTTGTTCTTAATTTCTGTGAGACTGTCTGGACCAAGCAACTCAG 2279
2221 ACAAAGCACAAGACAGATTTGTTCTTAATTTCTGTGAGACTGTCTGGACCAAGCAACTCAG 2280
2280 AAGTTTAAATATGCTCTCTTTGGATAAAATGAGTGGGTCAATGCTTTCAGAAAATGGAATC 2339
2281 AAGTTTAAATATGCTCTCTTTGGATAAAATGAGTGGGTCAATGCTTTCAGAAAATGGAATC 2340
2340 CAGTAGAATATGATACCAAGATGATCTTTGTAATTTAAATGCTGCTGATTCATCTCTCC 2399
2341 CAGTAGAATATGATACCAAGATGATCTTTGTAATTTAAATGCTGCTGATTCATCTCTCC 2400
2400 TGCTGAAAATGAAGCCACTGGCTTGGTCTTCTCCACCTCTTCTCCCAATCAGAGG 2459
2401 TGCTGAAAATGAAGCCACTGGCTTGGTCTTCTCCACCTCTTCTCCCAATCAGAGG 2460
2460 TCCATTGTTTCCAGTGGATGCAAGAGGCCCATTTCTTGAGAAAGAGACCTCTCTTCCCTCC 2519
2461 TCCATTGTTTCCAGTGGATGCAAGAGGCCCATTTCTTGAGAAAGAGACCTCTCTTCCCTCC 2520
2520 ACCTCTCCAGGACCATGTTTGGAGCTTCTCGAGATTTATTTCCACCAAGGATTTTCC 2579
2521 ACCTCTCCAGGACCATGTTTGGAGCTTCTCGAGATTTATTTCCACCAAGGATTTTCC 2580
2580 AGGTCACACACTGCTCCATTTGCAATGAGAAATGCTCTATCCACCGAGGGTCTTCTCTCC 2639
2581 AGGTCACACACTGCTCCATTTGCAATGAGAAATGCTCTATCCACCGAGGGTCTTCTCTCC 2640
2640 TTACCTTCCCAAGACCTGGAATTTTCCCTCCACCCACCCACATTTCTGAAGTAGAAGTGA 2699
2641 TTACCTTCCCAAGACCTGGAATTTTCCCTCCACCCACCCACATTTCTGAAGTAGAAGTGA 2700
2700 GTTCCCTCAGGTTGATTTCCACCTTCAATGAGGCTGCTACTGAACATCCAGAACCA 2759
2701 GTTCCCTCAGGTTGATTTCCACCTTCAATGAGGCTGCTACTGAACATCCAGAACCA 2760
2760 GCAAGAACCTGACAATATTTTGTCTCTTTCAAAAGTAAATTTGACTGATCTCAATTTT 2819
2761 GCAAGAACCTGACAATATTTTGTCTCTTTCAAAAGTAAATTTGACTGATCTCAATTTT 2820
2820 AGTTAAAGTAACTGCTGTTACTTAAAGTATACACTTTTGTCTCAAAATTTGAAGCTTAAATGG 2879
2821 AGTTAAAGTAACTGCTGTTACTTAAAGTATACACTTTTGTCTCAAAATTTGAAGCTTAAATGG 2880
2880 AATTATAATTTCTCAGGATAGTATTTTGTAAATPAAAGATGATTTAAATATGAATCTTATGA 2939
2881 AATTATAATTTCTCAGGATAGTATTTTGTAAATPAAAGATGATTTAAATATGAATCTTATGA 2940
2940 GTAATATTTTCAATTTTATTTAGACGCTATAACTATTTCAATTTGATTGATTATCCACTAT 2999
2941 GTAATATTTTCAATTTTATTTAGACGCTATAACTATTTCAATTTGATTGATTATCCACTAT 3000
3000 TATATAAACAAATAGTGGGAGTTTATATATATATATATATATATATATATATATATATAT 3059
3001 TATATAAACAAATAGTGGGAGTTTATATATATATATATATATATATATATATATATATAT 3060
3060 CTGAAGTCTGTCTTTATGCCAGAACTGATTTACTTGTGTTGTTGTTGTTGTTGTTGTTGTTG 3119
3061 CTGAAGTCTGTCTTTATGCCAGAACTGATTTACTTGTGTTGTTGTTGTTGTTGTTGTTGTTG 3120
3120 GTAACCTTTATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3164
3121 GTAACCTTTATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3165

PCT-US02-25465-51

; Sequence SI, Application PC/TUS0225465
; GENERAL INFORMATION:

; APPLICANT: INCYTE GENOMICS, INC.

; APPLICANT: AZIMZAI, Yalda

; APPLICANT: BARROSO, Ines

; APPLICANT: BAUGHN, Mariah R.

; APPLICANT: BECHA, Shanya D.

; APPLICANT: BOROWSKI, Mark L.

; APPLICANT: DUGGAN, Brendan M.

; APPLICANT: ELLIOTT, Vicki S.

; APPLICANT: EMERLING, Brooke M.

; APPLICANT: FORSYTHE, Ian J.

; APPLICANT: GIETZEN, Kimberly J.

; APPLICANT: GORVAD, Ann E.

; APPLICANT: GRIFFIN, Jennifer C.

; APPLICANT: GURURAJAN, Rajagopal

; APPLICANT: HAFALIA, April J.A.

; APPLICANT: ISON, Craig H.

; APPLICANT: KABLE, Amy E.

; APPLICANT: KHAN, Farrah A.

; APPLICANT: LEE, Sally

; APPLICANT: LEE, Soo Yeun

; APPLICANT: LI, Joana X.

; APPLICANT: REDDY, Roopa

; APPLICANT: RICHARDSON, Thomas W.

; APPLICANT: SPRAGUE, William W.

; APPLICANT: SWARNAKAR, Anita

; APPLICANT: TANG, Y. Tom

; APPLICANT: WARREN, Bridget A.

; APPLICANT: XU, Yuming

; APPLICANT: YAO, Monique G.

; APPLICANT: YUE, Henry

; APPLICANT: YUE, Huibin

; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH

; FILE REFERENCE: PP-1126 PCT

; CURRENT APPLICATION NUMBER: PCT/US02/25465

; CURRENT FILING DATE: 2002-08-08

; PRIOR APPLICATION NUMBER: US 60/311,017

; PRIOR FILING DATE: 2001-08-08

; PRIOR APPLICATION NUMBER: US 60/313,070

; PRIOR FILING DATE: 2001-08-17

; PRIOR APPLICATION NUMBER: US 60/313,071

; PRIOR FILING DATE: 2001-08-17

; PRIOR APPLICATION NUMBER: US 60/314,678

; PRIOR FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: US 60/316,692

; PRIOR FILING DATE: 2001-08-31

; PRIOR APPLICATION NUMBER: US 60/317,913

; PRIOR FILING DATE: 2001-09-07

; PRIOR APPLICATION NUMBER: US 60/322,182

; PRIOR FILING DATE: 2001-09-14

; PRIOR APPLICATION NUMBER: US 60/340,747

; PRIOR FILING DATE: 2001-12-07

; PRIOR APPLICATION NUMBER: US 60/342,761

; PRIOR FILING DATE: 2001-12-20

; PRIOR APPLICATION NUMBER: US 60/369,129

; PRIOR FILING DATE: 2002-03-29

; NUMBER OF SEQ ID NOS: 54

; SOFTWARE: PERL Program

; SEQ ID NO 51

; LENGTH: 3108

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; OTHER INFORMATION: Incyte ID No: 7506167CBI

PCT-US02-25465-51

Query Match

Best Local Similarity 91.0%; Score 2880.4;

Matches 3023; Conservative 0; Mismatches 6; Indels 129; Gaps 1;

QY 1 AGGTTTATCCATGAAGAAGACAGCAATTTTAAAAAGTGTTATTCACCAAAAAATAAGCTT 60
Db 72 AGGTTTATCCATGAAGAAGACAGCAATTTTAAAAAGTGTTATTCACCAAAAAATAAGCTT 131
QY 61 CAAATATGTGATGTGAAAACTGCCAGAACTAAAGCGCGCGCGCTCAGACAGCGCTGC 120
Db 132 CAAATATGTGATGTGAAAACTAAAGCGCGCGCGCTCAGACAGCGCTGC 191
QY 121 CTCAGGATGTAAAGTGTAAACAAGAGGCGAGGGAGGTGGTGGGGACAAATATGGCCCTG 180
Db 192 CTCAGGATGTAAAGTGTAAACAAGAGGCGAGGGAGGTGGTGGGGACAAATATGGCCCTG 251
QY 181 TGAGGCGCTGTGGTGGCGCCCGCTTCCCGAGCTCCCGCGAGCGCGCTCCACAGTGGTCCG 240
Db 252 TGAGGCGCTGTGGTGGCGCCCGCTTCCCGAGCTCCCGCGAGCGCGCTCCACAGTGGTCCG 311
QY 241 CTCGGTGTGGTGTACGTGGCGCATTCGGGTTCAGAGCCAAAGGCTGCGTGTCTCCACC 300
Db 312 CTCGGTGTGGTGTACGTGGCGCATTCGGGTTCAGAGCCAAAGGCTGCGTGTCTCCACC 371
QY 301 GCTTGTGTGGCGAGGTACTGCGGTGACCGCGAGCAGCGCTCGACCTATGGAGGAG 360
Db 372 GCTTGTGTGGCGAGGTACTGCGGTGACCGCGAGCAGCGCTCGACCTATGGAGGAG 431
QY 361 CCTGGTGTACCCCTCAGCCCTACCTGGGCTGGTCTCGAGGAGCTACGCGAGTGTG 420
Db 432 CCTGGTGTACCCCTCAGCCCTACCTGGGCTGGTCTCGAGGAGCTACGCGAGTGTG 491
QY 421 GCAGCACTACCTGAGAGTATGAGCCAGATGAGATCCCTTATGGTTTCCATCGGAACCTG 480
Db 492 GCAGCACTACCTGAGAGTATGAGCCAGATGAGATCCCTTATGGTTTCCATCGGAACCTG 551
QY 481 GTGGTATGTGCAGCTGTATTGGATTTTGTGTTTCTCTCTTTTGTGGAGAGTGT 540
Db 552 GTGGTATGTGCAGCTGTATTGGATTTTGTGTTTCTCTCTTTTGTGGAGAGTGT 611
QY 541 AGATCGGTAGGAGTGGCTTTTACGTGGGAGAGACGACAAACTTGGTCAACGCTTCT 600
Db 612 AGATCGGTAGGAGTGGCTTTTACGTGGGAGAGACGACAAACTTGGTCAACGCTTCT 671
QY 601 GGAATAATGAAGAAAAATTAATACTACTTTGAAAAATTTAGCCTTATTCAAAAAGAGTAT 660
Db 672 GGAATAATGAAGAAAAATTAATACTACTTTGAAAAATTTAGCCTTATTCAAAAAGAGTAT 731
QY 661 GAAGGCTATGAAGTAGATCATCTTTAGAGATGCGAGCTTTGAGAGGAGGCGAGCAGAA 720
Db 732 GAAGGCTATGAAGTAGATCATCTTTAGAGATGCGAGCTTTGAGAGGAGGCGAGCAGAA 791
QY 721 GAAGCAAGAAAGTTGGAGGCACTCTGAAAAAGCTGAAACAGGTCCCAATTCGAACTTGAG 780
Db 792 GAAGCAAGAAAGTTGGAGGCACTCTGAAAAAGCTGAAACAGGTCCCAATTCGAACTTGAG 851
QY 781 GATGAATTCCTCTGTCTAGAAAAAGACTTAAACAAGAGAAATCTAAACATTTCTCAACAA 840
Db 852 GATGAATTCCTCTGTCTAGAAAAAGACTTAAACAAGAGAAATCTAAACATTTCTCAACAA 911
QY 841 GATGAATTCCTCTGTCTAGAAAAAGACTTAAACAAGAGAAATCTAAACATTTCTCAACAA 900
Db 912 GATGAATTCCTCTGTCTAGAAAAAGACTTAAACAAGAGAAATCTAAACATTTCTCAACAA 971
QY 901 CTCAAATCACAATAGCTGAAGCAAAATCATCTGCAAGACATTTAAATGAGTGAAGAA 960
Db 972 CTCAAATCACAATAGCTGAAGCAAAATCATCTGCAAGACATTTAAATGAGTGAAGAA 1031
QY 961 CGACGGCTATAGCAATAAAGATGCTTTCGAATGAAATTTCTCAACTTCAGACAAAGCAT 1020
Db 1032 CGACGGCTATAGCAATAAAGATGCTTTCGAATGAAATTTCTCAACTTCAGACAAAGCAT 1091
QY 1021 AAACAGCTTTTTCAGCAAGAGCTGAAAGTATGAAAGGAGAGAGTGAAGTAAATAAA 1080
Db 1092 AAACAGCTTTTTCAGCAAGAGCTGAAAGTATGAAAGGAGAGAGTGAAGTAAATAAA 1151

QY 1081 CAGAAAATAACATTGTAAGACTCCAAAGTACACGACAAAGCTTCTGATGATTAAGAA 1140
Db 1152 CAGAAAATAACATTGTAAGACTCCAAAGTACACGACAAAGCTTCTGATGATTAAGAA 1211
QY 1141 AATACATCAAGACCCCTGACTGACACTCTCCAAATGATGAAGATCAGGCTGTGCTT 1200
Db 1212 AATACATCAAGACCCCTGACTGACACTCTCCAAATGATGAAGATCAGGCTGTGCTT 1271
QY 1201 GAAGAAGACACAACGAGTATGATTAACCTCGAATTTAGAGTGAACGATGAATCGGAAT 1260
Db 1272 GAAGAAGACACAACGAGTATGATTAACCTCGAATTTAGAGTGAACGATGAATCGGAAT 1331
QY 1261 GGTGCTTACTTAGATAATCTCCAAAAGGAGCTTTGAAGAACTGATTTATGCTGCTAAG 1320
Db 1332 GGTGCTTACTTAGATAATCTCCAAAAGGAGCTTTGAAGAACTGATTTATGCTGCTAAG 1391
QY 1321 TTAATGCTTCTTTAAAACCTTAGAGGAGAAAGAACTTTAATATTTATATTCAGTTGCT 1380
Db 1392 TTAATGCTTCTTTAAAACCTTAGAGGAGAAAGAACTTTAATATTTATATTCAGTTGCT 1451
QY 1381 GAAGTTGATATAAACAAGGAGAGCTTACAGAGCATATTTAAAATCTTCAGACTCAACAA 1440
Db 1452 GAAGTTGATATAAACAAGGAGAGCTTACAGAGCATATTTAAAATCTTCAGACTCAACAA 1511
QY 1441 GCATCTTGAGTACAGTACAGAAACACACATTTTGAATAATGAGAACTCAGAGCTTCAACAGAAA 1500
Db 1512 GCATCTTGAGTACAGTACAGAAACACACATTTTGAATAATGAGAACTCAGAGCTTCAACAGAAA 1571
QY 1501 CTTAAAGTAACTGACTGAATTTATCAAGAAATGAAATGAACTCCACAGGAAATTAACA 1560
Db 1572 CTTAAAGTAACTGACTGAATTTATCAAGAAATGAAATGAACTCCACAGGAAATTAACA 1631
QY 1561 GTAGAGAAAATTTATCGGTTAGAGAAAGAGAACTTTCTAAAGTAGAGTGAAGATC 1620
Db 1632 GTAGAGAAAATTTATCGGTTAGAGAAAGAGAACTTTCTAAAGTAGAGTGAAGATC 1691
QY 1621 AGCCATCCACTGAAGAGCTGGAGCTATAGAAAGCGAGCCAAAGATCTTGAGAGAA 1680
Db 1692 AGCCATCCACTGAAGAGCTGGAGCTATAGAAAGCGAGCCAAAGATCTTGAGAGAA 1751
QY 1681 TTGAGAGAACTATTCATTTCTTAAGGCGAGATTTATTTCCATGAGAAAAAGCACAT 1740
Db 1752 TTGAGAGAACTATTCATTTCTTAAGGCGAGATTTATTTCCATGAGAAAAAGCACAT 1811
QY 1741 GATAATTTGGTGGAGCTCGGATCTGGAAGAACTCAATGATTTAGAGGAGAAAT 1800
Db 1812 GATAATTTGGTGGAGCTCGGATCTGGAAGAACTCAATGATTTAGAGGAGAAAT 1871
QY 1801 GCTCAACACAGACAAAAATTAACCTGAAACAGAGCTTAAATTTGAACTTTTAGAAAAAGAT 1860
Db 1872 GCTCAACACAGACAAAAATTAACCTGAAACAGAGCTTAAATTTGAACTTTTAGAAAAAGAT 1931
QY 1861 CCTATGCACTCGATGTTCCAAATACAGCATTTGGCAGAGAGCATTTCCCATATGTCCTCC 1920
Db 1932 CCTATGCACTCGATGTTCCAAATACAGCATTTGGCAGAGAGCATTTCCCATATGTCCTCC 1967
QY 1921 TCACCATTTGGGTTGGCTTCTATCTGAAACAGAGCTTTTCTCTCTCTCCCACTTTGTTG 1980
Db 1968 ----- 1967
QY 1981 GAGGTCACCTCAGACTCTCACCTTTCCAGGGGAGAGAGAGGCTCAGAGGC 2040
Db 1968 -----AGAGGCTCAGAGGC 1982
QY 2041 CCAGGAACTCTCTGGACCTCAGATTTACCAATGAAAGAGAGAACTCAAGCTCTGATAGG 2100
Db 1983 CCAGGAACTCTCTGGACCTCAGATTTACCAATGAAAGAGAGAACTCAAGCTCTGATAGG 2042
QY 2101 TTAACCGATCTCATAGGCTCCCTCTGACATGGGCTCTGTCACTCCATCGGACCGAG 2160
Db 2043 TTAACCGATCTCATAGGCTCCCTCTGACATGGGCTCTGTCACTCCATCGGACCGAG 2102
QY 2161 GACCGTAGGATGATTTCTCCCGCCAGGACAAATCATATCTGATTCAGGCCCTCTCCCA 2220

Db 2103 GACCGTAGGATGATTTCTCCCGCCAGGACAAATCATATCTGATTCAGCCCTCTCTCCA 2162
QY 2221 CAAAGGCAAGACAGATTTTGTCTTAATTTCTGGTAGACTGTCTGGACCAAGAGAACTCAGA 2280
Db 2163 CAAAGGCAAGACAGATTTTGTCTTAATTTCTGGTAGACTGTCTGGACCAAGAGAACTCAGA 2222
QY 2281 AGTTTAAATATGCTCTTTCTGGATAAAATGATGGGTCAATGCCCTTCAGAAATGGAATCC 2340
Db 2223 AGTTTAAATATGCTCTTTCTGGATAAAATGATGGGTCAATGCCCTTCAGAAATGGAATCC 2282
QY 2341 AGTAGAAATGATACAAAGATGATCTTTGGTAAATTTAAATGTGCCGTATCATCTCCCT 2400
Db 2283 AGTAGAAATGATACAAAGATGATCTTTGGTAAATTTAAATGTGCCGTATCATCTCCCT 2342
QY 2401 GCTGAAATGAGCCACACCTGGCCCTTCTCTCCACCTCTTGTCTCCAAATCAGAGGT 2460
Db 2343 GCTGAAATGAGCCACACCTGGCCCTTGTCTCTCCACCTCTTGTCTCCAAATCAGAGGT 2402
QY 2461 CCATTTGTTTCCAGTGGATGCAAGAGGCCCATTTCTTGAGAGAGAGGACCTCTTTCCCCCA 2520
Db 2403 CCATTTGTTTCCAGTGGATGCAAGAGGCCCATTTCTTGAGAGAGGACCTCTTTCCCCCA 2462
QY 2521 CTTCTCTCAGAGGACCATGTTTGGAGCTTCTCGAGATTTATTTTCCACCAAGAGGATTTCCCA 2580
Db 2463 CTTCTCTCAGAGGACCATGTTTGGAGCTTCTCGAGATTTATTTTCCACCAAGAGGATTTCCCA 2522
QY 2581 GGTCCACACACTGTCTCCATTTTGAATGAGAAATGTCTATCCACGAGGGGTTTCTCTCT 2640
Db 2523 GGTCCACACACTGTCTCCATTTTGAATGAGAAATGTCTATCCACGAGGGGTTTCTCTCT 2582
QY 2641 TACCTTCCCCCAAGACCTGGATTTTCCCCCAAGGACCATTTCTGAAGTAGAGTAGAG 2700
Db 2583 TACCTTCCCCCAAGACCTGGATTTTCCCCCAAGGACCATTTCTGAAGTAGAGTAGAG 2642
QY 2701 TTCCCTCTCAGGTTTGAATTTCCACCTTCAATGAGCTGTCTGAACTCCAGAACCCAG 2760
Db 2643 TTCCCTCTCAGGTTTGAATTTCCACCTTCAATGAGCTGTCTGAACTCCAGAACCCAG 2702
QY 2761 CAAAGAACTGACAAATATTTTGTCTCTCTTCAAAAGTAAATTTTGAATCTCATTTTCA 2820
Db 2703 CAAAGAACTGACAAATATTTTGTCTCTCTTCAAAAGTAAATTTTGAATCTCATTTTCA 2762
QY 2821 GTTTAAAGTAACTGCTGTACTTAAAGTAACTTAAAGTAAATTTGAATCTCATTTTCA 2880
Db 2763 GTTTAAAGTAACTGCTGTACTTAAAGTAACTTAAAGTAAATTTGAATCTCATTTTCA 2822
QY 2881 ATTATAATTTCTCAGGATAGTATTTTGAATTAAGATGATTTAAATATGAATCTTATGAG 2940
Db 2823 ATTATAATTTCTCAGGATAGTATTTTGAATTAAGATGATTTAAATATGAATCTTATGAG 2882
QY 2941 TAAATTTTCTCAATTTTATTTTAGCGGTATACTATTTCAATTTGATTAATCCACTATT 3000
Db 2883 TAAATTTTCTCAATTTTATTTTAGCGGTATACTATTTCAATTTGATTAATCCACTATT 2942
QY 3001 ATATAAACTAGTGGGAGTTTATATATGTAATTTCTTTCAGGTGGGAGGCTTTAAATTC 3060
Db 2943 ATATAAACTAGTGGGAGTTTATATATGTAATTTCTTTCAGGTGGGAGGCTTTAAATTC 3002
QY 3061 TGAAGTCTGTCTTTTATGCGCAAGACCTGATTTTACTGTGGTTGTGGCAATCTGGAAG 3120
Db 3003 TGAAGTCTGTCTTTTATGCGCAAGACCTGATTTTACTGTGGTTGTGGCAATCTGGAAG 3062
QY 3121 TAACTTTTATGCTTAAATAAATTTATGTTGATTTAAAAA 3158
Db 3063 TAACTTTTATGCTTAAATAAATTTATGTTGATTTAAAGA 3100

RESULT 9

US-60-500-315-11859/c

; Sequence 11859, Application US/60500315

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele

Qy	1861	CCTTATGCACTCGATGTTCCAAATACAGCATTTGGCAGAGAGCATTCGCCATATGGTCCC	1920
Db	78279	CCTAATGCACTCGATGTTTCAAAATACAGCATTTGGCAGAGAGCATTCGCCATATGGTCCC	78220
Qy	1921	TCACCATTTGGTTTGGCCTTCATCTGAAAACAAGAGCTTTTCTCTCTCCCTCCAACTTTGTTG	1980
Db	78219	TCACCATTTGGTTCGGCTTCATCTGAAAACGAGAGCTTTTCCCTCTCTCCAAACTTTGTTG	78160
Qy	1981	GAGGTCCTCATAGACTCTCACCTTTTCTGTTTCCAGGGGGAGGAGGAGGCTCACGAGC	2040
Db	78159	GAGGATCCACTCAGACTCTCACCTGTCGTTTCCAGGGGGAGGAGGAGGCTCCAAAGCAGC	78100
Qy	2041	CCAGGGAATCCTCTGGACCATCAGATTACCAATGAAAGAGGAGAGATCAAGCTGTGATAGG	2100
Db	78099	CCAGGGAATCCTCTGGACCATCAGATTACCAATGAAAGAGGAGAGATCAAGCTGTGACAGG	78040
Qy	2101	TTAACCGATCCTCATAGGCTCCTCTGACACTGGGTCTCTGTCACTCTCCATGGGACCCAG	2160
Db	78039	TTAATCGATCCTCACAGGCTCCTCTGTGACACTGGGTCTCTGTCACTCTCCGCTGGAACAG	77980
Qy	2161	GACCGTAGGATGATGTTTCTCTCCGCGCAGGACCAATCATATCCTTGATTCAGCCCTTCTCCA	2220
Db	77979	GACCGTAGGATGATGTTTCTCTCACCAGGCAATCATATCCTTGATTCAACTCTTCTCTCCA	77920
Qy	2221	CAAAGGCAACACAGATTTTGTCTTAATTTCTGTAGACTCTCTGGACCCAGCAGAACTCAGA	2280
Db	77919	CAAAGGCAACACAGATTTTATTTCTAATTTCTGAAAGACTCTCTGGACCCAGCAGAACTCAGA	77860
Qy	2281	AGTTTAAATATGCTTCTTTTGGATAAAATGGATGGGTCAATGCTTTCAGAAATGGAATCC	2340
Db	77859	AGTTTAAATATGCTTCTTTTGGATAAAATGGATGGGTCAATGCTTTCAGAAATGGAATCC	77800
Qy	2341	AGTAGAAATGATACCAAAGATGATCTTTGGTAAATTTAAATGTGCTTGATTCATCTCTCCCT	2400
Db	77799	AGTAGAAATGATGCCAAGATGATCTTTGGTAAATTTAAATGTGCTTGATTCATCTCTCCCT	77740
Qy	2401	GCTGAAATGAAGCACTGGCCCTGGCTTTTTCCTCCACCTCTTGCTCCAAATCAGAGGT	2460
Db	77739	GCTGAAATGAAGCACTGGCCCTGGCCCTTATTCCTCYACCTCTTGCTCCAGTCAGCGGT	77680
Qy	2461	CCATTGTTCCAGTGGATGCAAGAGGCCCACTCTTCAGAAAGAGGACCTCTTTTCCCCCA	2520
Db	77679	CCATTGTTCCAGTGGATGCAAGAGGCCCTTCTAGAAAGAGGACCTCTTTTCCCCCA	77620
Qy	2521	CCTCTCCAGGACCAATGTTTGGAGCTTCTCGAGATTTATTTTCCACCAAGGATTTCCCA	2580
Db	77619	CCTCTCCAGGACCAATGTTTGGAGCTTCTCGAGGTTATTTTCCCAAGGATTTTCCCA	77560
Qy	2581	GGTCCACACCTGCTCCATTGCAATGAGNAATGCTATTCACCGAGGGGTTTCTCTCTCT	2640
Db	77559	GGTCCACACATGCTCCATTGCAATGAGNAACATCTATCCACCGAGGGGTTTCTCTCTCT	77500
Qy	2641	TACCTTCCCCCAAGACCTGGATTTTTCCTCCCAACCCCACTTCTGAAGGTAGAGTAGAG	2700
Db	77499	TACCTTCAACCGAGACCTGGATTTT - ACCCCAAACCCACATCTGAAGGTAGAGCSAG	77442
Qy	2701	TTCCCTCCAGGTTTGATTCACCTTCAATGAGCCCTGCTACTGAAATCCAGAACACAG	2760
Db	77441	TTCCCTTTCAGGATTTGATTCGCGCTTCAAGAGGACCTGCTACTGGAATCCAGAACACAG	77382
Qy	2761	CAAGAAACCTGACAAATATTTTGTCTCTCTCAAAAGTAAATTTTGTACTGATCTCATTTTCA	2820
Db	77381	CAAGAAACCTGACAAATATTTTGTCTCTCTCAAAAGTAAATTTTGTACTGATCTCATTTTCA	77322
Qy	2821	GTTTAAAGTAATGCTGTATCTTAAGTGATTTACACTTTTCTCTCAAAATGGAAGCTTAATGA	2880
Db	77321	GTTTAAAGTAATGCTGTATCTTAAGTGATTTACACTTTT - CTCAAAATGGAAGCTTAATGA	77263
Qy	2881	ATTATATTTCTCAGGATAGATTTTGTGAAATAAGATGATTTTAAATGAATCTTATGAG	2940
Db	77262	ATAATAGTTCTCAGGATAGATTTTGTGAAATAAGATGTTTGAATGATGAATCTTATGAG	77203
Qy	2941	TAAATTTTCAATTTT - - - - - TATTTTAGACGGTAACTATTTTCAAT - - - - - TTGATTAATC	2993

[illegible]

Db 523 TTTATGTGGACGAGAGAAAGAGCTTCTCTAATCTCTTCGACTAATTGAGAAATA 582
QY 620 GTAACTACTTGAATAATTTAGCTTATTTCAAAAAGAGTATGAAGCTATGAAGTAGT 679
Db 583 GTAACTACTTGAATAATTTAGCTTATTTCAAAAAGAGTATGAAGCTATGAAGTAGT 642
QY 680 CATCTTTAGAGTATCCAGCTTTTGAGAGGAGCCAGAGAAAGCAGAAAGTATGGAGG 739
Db 643 CATCTTTAAGGATGCCAGCTTTTGAGAGGAGCCAGAGAAAGCAGAAAGTATGGAGG 699
QY 740 CAACCTGTGAAGAGTGAACAGCTTCAATCTGAATCTTGAGGATGAATTCCTGTCTAG 799
Db 700 CAACCTGTGAAGAGTGAACAGCTTCAATCTGAATCTTGAGGATGAATTCCTGTCTAG 759
QY 800 AAAAAAGCTTTAAAAAAGAGAGAAATCTAAACATCTCTCAACAGATCAATTTGATGGCGATA 859
Db 760 AAAAAAGCTTTAAAAAAGAGAGAAATCCAAACATCTCTGAACAGATCAATTTGATGGCGATA 819
QY 860 TTTCAAAAGTATACAGTCTCTAGAGATGATGATCAAAATCCCTCAAAATCAAAATAGCTG 919
Db 820 TTTCAAAAGGATACAGTCTCTAGAGATGATGATCAAAATCCCTCAAAATCAAAATAGCTG 879
QY 920 AAGCCAAATCATCTGCAAGACATTTAAATCAGTGAAGAGCGAGCGCTATAGCAATAA 979
Db 880 AAGCCAAATGACCTTCAAGATATTTCAATGAATGAAGAGCGACTGAAGATAGCAATAA 939
QY 980 AAGATGCTTTGAATGAATAATCTCACTTCAGACAAAGCATAAACAGCTTTTTCAGCAAG 1039
Db 940 AAGATGCTTTGAATGAATAATCTCACTTCAGACAAAGCGAGAGAGAGAGCTTTTTCAGCAAG 996
QY 1040 AAGCTGAAGTATGGAAGAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1099
Db 997 AAGCTGAAGTATGGAAGAGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1056
QY 1100 ACTCAAAGTACCGCAGAACCAAGTCTGAATGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1159
Db 1057 ACTCAAAGTACCGCAGAACCAAGTCTGAATGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1116
QY 1160 CTGGACATCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1219
Db 1117 CTGAAGCTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1176
QY 1220 ATGATAACCTGGAATTTAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1279
Db 1177 ATGATAACCTGGAATTTAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1236
QY 1280 CTCCAAAGAGCTTTGAAGAACCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1339
Db 1237 CTCCAAAGAGCTTTGAAGAACCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1296
QY 1340 CCTTGAAGAGGAGAAAGAAACCAAAATTTATATTTAGTGTCTGAAGTGAAGTGAAGTGAAGTGAAG 1399
Db 1297 CCTTGAAGAGGAGAAAGAACCAAAATTTATATTTAGTGTCTGAAGTGAAGTGAAGTGAAGTGAAG 1356
QY 1400 AAGAGCTTACAGAGCATTTAAATTTCTTCAGCTCAACAGCATCTTTGAGTGAAGTGAAGTGAAG 1459
Db 1357 AAGAGCTTACAGAGCATTTAAATTTCTTCAGCTCAACAGCATCTTTGAGTGAAGTGAAGTGAAG 1416
QY 1460 ACACACATTTGAAAATGAGATCAGAGCTTCAACAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1519
Db 1417 ACACACATTTGAAAATGAGATCAGAGCTTCAACAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1476
QY 1520 TATATCAAGAAATGAATGAATCTTCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1579
Db 1477 TATATCAAGAAATGAATGAATCTTCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1536
QY 1580 TAGAGAAAGAGAGAACTTTCTTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1639
Db 1537 TAGAGAAAGAGAGAACTTTCTTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1596
QY 1640 TGAGACCTTATAGAAAGCGAGCAAGATCTTTGAAGAGAAATTTGAGAGAACTATTTCTATT 1699

Db 1597 TGGAGACCTATAGAAAGCGAGCCAAAGATCTTTGAAGAGAAATTTGGAGAGAACTATTCTATT 1656
QY 1700 CTTATCAAGGGAGATATTTTCCCATGAGAGAAAGACATGATATAATTTGGTTGGCAGCTC 1759
Db 1657 CTTATCAAGGGAGATATTTTCCCATGAGAGAAAGACATGATATAATTTGGTTGGCAGCTC 1716
QY 1760 GGAATGCTGAAGAAACCTCAATGATTTAAGGAAAGAAATGCTCAACACAGACAAATAAT 1819
Db 1717 GGAATGCTGAAGAAACCTCAATGATTTAAGGAAAGAAATGCTCAACACAGACAAATAAT 1776
QY 1820 TAACTGAACAGAGCTTTAAATTTGAACTTTTGAAGAAAGATCTTTATGACTGATGTTTC 1879
Db 1777 TAACTGAACAGAGCTTTAAATTTGAACTTTTGAAGAAAGATCTTTATGACTGATGTTTC 1836
QY 1880 CAAATACAGATTTGCGAGAGCATTTCCCATATGCTCCCTCACCATTGGGTTGGCCTT 1939
Db 1837 CAAATACAGATTTGCGAGAGCATTTCCCATATGCTCCCTCACCATTGGGTTGGCCTT 1896
QY 1940 CATCTGAACAGAGCTTTTCTCTCTCTCCCTCAACTTTTGTGGAGGTCCTCAGACTCT 1999
Db 1897 CATCTGAACAGAGCTTTTCTCTCTCTCCCTCAACTTTTGTGGAGGTCCTCAGACTCT 1956
QY 2000 CACCTTTGCTTCCAGGGGAGAGAGAGCTTCAAGAGGCTTCAAGAGGCTTCAAGAGGCTTCAAGAGG 2059
Db 1957 CACCTTTGCTTCCAGGGGAGAGAGAGCTTCAAGAGGCTTCAAGAGGCTTCAAGAGGCTTCAAGAGG 2016
QY 2060 ATCAGATTTACCAATGAAGAGAGAGAAATCAAGCTGTGATAGGTTAAACCGATCTCTATAGG 2119
Db 2017 ATCAGATTTACCAATGAAGAGAGAGAAATCAAGCTGTGATAGGTTAAACCGATCTCTATAGG 2076
QY 2120 CTCCCTCTGACATGCTGTCTCTACCTTCCATGGGACAGAGGCTTCAAGAGGCTTCAAGAGGCTTCAAGAGG 2179
Db 2077 CTCCCTCTGACATGCTGTCTCTACCTTCCATGGGACAGAGGCTTCAAGAGGCTTCAAGAGGCTTCAAGAGG 2136
QY 2180 CTCGGCAGAGCAATCATATCTCTGATTCAGCCCTTCTCTCCAAAGGCAAGACAGATTTT 2239
Db 2137 CTCGGCAGAGCAATCATATCTCTGATTCAGCCCTTCTCTCCAAAGGCAAGACAGATTTT 2196
QY 2240 GTTCTAATCTGTAGATGCTGTGGACAGAGCAACTCAGAGTCTCAGAGTCTTAAATATGCTCTTCT 2299
Db 2197 GTTCTAATCTGTAGATGCTGTGGACAGAGCAACTCAGAGTCTTAAATATGCTCTTCTTCTTCT 2256
QY 2300 TGGATAAATGATGGGTCAATGCTTTCAGAAATGGAATCCAGTGAAGTGAAGTGAAGTGAAGTGAAG 2359
Db 2257 TGGATAAATGATGGGTCAATGCTTTCAGAAATGGAATCCAGTGAAGTGAAGTGAAGTGAAGTGAAG 2316
QY 2360 ATGATCTGTGTAATTTAAATGCTGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2419
Db 2317 ATGATCTGTGTAATTTAAATGCTGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2376
QY 2420 GCCTTGGCTTTGCTCTCCACCTTCTGCTTCAATCAGAGTCTCATTGTTTCCAGTGGATG 2479
Db 2377 GCCTTGGCTTTGCTCTCCACCTTCTGCTTCAATCAGAGTCTCATTGTTTCCAGTGGATG 2436
QY 2480 CAAAGGCCCATTTCTTGAGAGAGAGCTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2539
Db 2437 CAAAGGCCCATTTCTTGAGAGAGAGCTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2496
QY 2540 TTGAGGCTTCTCAGATTTATTTTCCACCAAGGATTTCCAGGTCACACCTGCTCTCCAT 2599
Db 2497 TTGAGGCTTCTCAGATTTATTTTCCACCAAGGATTTCCAGGTCACACCTGCTCTCCAT 2556
QY 2600 TTGCAATGAGAAATGCTATCCACCGAGGGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2659
Db 2557 TTGCAATGAGAAATGCTATCCACCGAGGGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2616
QY 2660 GATTTTTTCCCCCCCCCAGATTTCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2719
Db 2617 GATTTTTTCCCCCCCCCAGATTTCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2676
QY 2720 CACCTTCAAAATGAGCTGTCTGATCAATCCAGAAACCAAGAGAGAAACCTGACAAATTT 2779
Db 2677 CACCTTCAAAATGAGCTGTCTGATCAATCCAGAAACCAAGAGAGAAACCTGACAAATTT 2736

1520 TATATCAAGAAATGAATGAACTCCACAGAAATTAACAGTAGAGGAAATTTATCGGT 1579
1541 TATATCAAGAAATGAATGAACTCCACAGAAATTAACAGTAGAGGAAATTTATCGGT 1600
1580 TAGAGAAAGAGAGAACTTTCTAAAGTAGATGAAAGATCAGCCATGCCACTGAGAGC 1639
1601 TAGAGAAAGAGAGAACTTTCTAAAGTAGATGAAAGATCAGCCATGCCACTGAGAGC 1660
1640 TGGAGACTATAGAAAGCGAGCAAGATCTTGAAGAGAAATGGAGAGAACTTCAAT 1699
1661 TGGAGACTATAGAAAGCGAGCAAGATCTTGAAGAGAAATGGAGAGAACTTCAAT 1720
1700 CTTATCAAGGGCAGATTAATTTCCCATGAGAAAAAGACATGATTAATTTGGTGGCAGCT 1759
1721 CTTATCAAGGGCAGATTAATTTCCCATGAGAAAAAGACATGATTAATTTGGTGGCAGCT 1780
1760 GGAATGCTGAAGAAACCTCAATGATTTAAGAAAGAAATGCTCACAACAGACAAAAT 1819
1781 GGAATGCTGAAGAAACCTCAATGATTTAAGAAAGAAATGCTCACAACAGACAAAAT 1840
1820 TAACTGAACAGAGCTTAAATTTGAACCTTTTGAAGAAAGATCCTTATGCACTCCGATGTC 1879
1841 TAACTGAACAGAGCTTAAATTTGAACCTTTTGAAGAAAGATCCTTATGCACTCCGATGTC 1900
1880 CAAATACAGATTTGGCAGAGAGCAATCCCATATGCTCCCTCACATTTGGGCTT 1939
1901 CAAATACAGATTTGGCAGAGAGCAATCCCATATGCTCCCTCACATTTGGGCTT 1960
1940 CATCTGAACAGAGCTTTCTCTCTCCCTCAACTTTTGTGGAGGCTCCACTCAGACCT 1999
1961 CATCTGAACAGAGCTTTCTCTCTCCCTCAACTTTTGTGGAGGCTCCACTCAGACCT 2020
2000 CACCTTTGCTCCAGGGGAGAGAGAGGCTCCAGAGGCCAGGAAATCCTCTGAGCC 2059
2021 CACCTTTGCTCCAGGGGAGAGAGAGGCTCCAGAGGCCAGGAAATCCTCTGAGCC 2080
2060 ATCAGATTTACCAATGAAGAGAGAAATCAAGCTGTGATAGGTTAAACCATCCTCATAGG 2119
2081 ATCAGATTTACCAATGAAGAGAGAAATCAAGCTGTGATAGGTTAAACCATCCTCATAGG 2140
2120 CTCCTCTGACACTGGTCTCTGTACCTCCATGGGACAGGACCGTAGGATGTTTC 2179
2141 CTCCTCTGACACTGGTCTCTGTACCTCCATGGGACAGGACCGTAGGATGTTTC 2200
2180 CTCGGCCAGACATCATATCTGTATCTAGCCCTTCCCTCCAAAGGCAAGACAGATTT 2239
2201 CTCGGCCAGACATCATATCTGTATCTAGCCCTTCCCTCCAAAGGCAAGACAGATTT 2260
2240 GTTCTAATCTGGTAGACTGTCTGGACAGCAGAACTCAGAAAGTTTAAATATGCTTCTT 2299
2261 GTTCTAATCTGGTAGACTGTCTGGACAGCAGAACTCAGAAAGTTTAAATATGCTTCTT 2320
2300 TGGATAAATGATGGTCAATGCTTCAAGATGGAATCCAGTAAATGATACCAAG 2359
2321 TGGATAAATGATGGTCAATGCTTCAAGATGGAATCCAGTAAATGATACCAAG 2380
2360 ATGATCTTGGTAAATTTAAATGTGCTGTATCTCTCCCTGCTGAAATGAAGCACTG 2419
2381 ATGATCTTGGTAAATTTAAATGTGCTGTATCTCTCCCTGCTGAAATGAAGCACTG 2440
2420 GCCTGGCTTTGTTCTCCACCTTTGCTCCAAATCAGAGGTCATTTTTCAGTGGATG 2479
2441 GCCTGGCTTTGTTCTCCACCTTTGCTCCAAATCAGAGGTCATTTTTCAGTGGATG 2500
2480 CAAAGGCCCATTCTTTGAGAGAGGACCTCTTTTCCCTCCCTCCTCAGGAGCCATG 2539
2501 CAAAGGCCCATTCTTTGAGAGAGGACCTCTTTTCCCTCCCTCCTCAGGAGCCATG 2560
2540 TTGGAGCTTCTCGAGATTAATTTTCCACCAAGGATTTCCAGGTCACACCTGCTCCAT 2599
2561 TTGGAGCTTCTCGAGATTAATTTTCCACCAAGGATTTTCCAGGTCACACCTGCTCCAT 2620

2600 TTGCAATGAGAAATGCTATCCACGAGGGGTTTTCTCTTACCTTCCCCCAAGACCTG 2659
2621 TTGCAATGAGAAATGCTATCCACGAGGGGTTTTCTCTTACCTTCCCCCAAGACCTG 2680
2660 GATTTTTCCCCCCCCACACATCTTGAAGGTAGAGTCCCTCAGGTTGATTC 2719
2681 GATTTTTCCCCCCCCACACATCTTGAAGGTAGAGTCCCTCAGGTTGATTC 2740
2720 CACCTTCAATGAGCCTGCTACTGACATCCAGAACCCAGCAAGAACTCAGCAATAT 2779
2741 CACCTTCAATGAGCCTGCTACTGACATCCAGAACCCAGCAAGAACTCAGCAATAT 2800
2780 TTTGCTCTCTTCAAAAGTAATTTTGAATGATCTTCAATTTTCAAGTAAGTCTGTTA 2839
2801 TTTGCTCTCTTCAAAAGTAATTTTGAATGATCTTCAATTTTCAAGTAAGTCTGTTA 2860
2840 CTTTARGTATACATTTTCTCAATTTGAAGTAAATGGAATTAATTTCTCAGGATAG 2899
2861 CTTTARGTATACATTTTCTCAATTTGAAGTAAATGGAATTAATTTCTCAGGATAG 2920
2900 TATTTTGAATTAAGATGATTTTAAATATGAATCTTATGAGTAAATTTTCAATTTTAT 2959
2921 TATTTTGAATTAAGATGATTTTAAATATGAATCTTATGAGTAAATTTTCAATTTTAT 2980
2960 TTTAGAGGTATACATTTTCAATTTTGAATTAATCCACTATTAATAACAATAGTGGAG 3019
2981 TTTAGAGGTATACATTTTCAATTTTGAATTAATCCACTATTAATAACAATAGTGGAG 3040
3020 TTTTATATATGTAATCTTTTCAAGTGGGAGGCTTTTAAATTTTCAAGTCTGTTTATG 3079
3041 TTTTATATATGTAATCTTTTCAAGTGGGAGGCTTTTAAATTTTCAAGTCTGTTTATG 3100
3080 CCAAGAACTGTAATTTTACTGTTTGGCAATGGAATGGAAGTAACTTTATGCTTAATAA 3139
3101 CCAAGAACTGTAATTTTACTGTTTGGCAATGGAATGGAAGTAACTTTATGCTTAATAA 3160
3140 ATTATAGTTGATTTAAAAA 3158
3161 ATTATAGTTGATTTAAGA 3179

RESULT 12
US-09-770-175-6675
; Sequence 6675, Application US/09770175
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.2058-001
; CURRENT APPLICATION NUMBER: US/09/770,175
; CURRENT FILING DATE: 2001-01-26
; PRIOR FILING DATE: 2000-01-28
; NUMBER OF SEQ ID NOS: 8967
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6675
; LENGTH: 3910
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-770-175-6675

Query Match 87.7%; Score 2773.8; DB 32; Length 3910;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 2927; Conservative 0; Mismatches 122; Indels 30; Gaps 5;

90 CTAAGGCGGCGGCTCAGACCCAGCGCTCCCTCAGATGTAAGTGAACAGAGGCG 149
107 CTTGTCGGGTCCGGTTCGGACCTCGCTCGGATGTAAGTGAACAGAGGCTC 166
150 AGGAGGAGTGGTGGGGACAAATGGCCCTGTGAGCCCTGTGGGTCCCGCTTCC 209
167 GGGATGGG-----CAGCGTAGGCTGTGAGGCTCGCGGTCGCCCTCTCC 216

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2004, 05:57:51 ; Search time 276 Seconds

(without alignments)
7705.114 Million cell updates/sec

Title: US-09-744-314-14

Perfect score: 3164

Sequence: 1 aggttcaatccatgaagaag.....agtgatttataaaaaaaa 3164

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1375775 seqs, 336063760 residues

Total number of hits satisfying chosen parameters: 2751550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA.New.*

1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2880.4	91.0	3108	US-10-486-020-51	Sequence 51, Appl
2	2515.2	79.5	2928	US-10-486-020-38	Sequence 38, Appl
3	201	6.4	260	US-10-793-479-29791	Sequence 29791, A
4	135.6	4.3	251	US-10-793-479-2603	Sequence 2603, Ap
5	133.6	4.2	140	US-10-793-479-27286	Sequence 27286, A
6	129.2	4.1	14393	US-10-021-698A-3524	Sequence 3524, Ap
7	64.8	2.0	265	US-10-793-479-34761	Sequence 34761, A
8	60.2	1.9	141646	US-60-550-051-3063	Sequence 3063, Ap
9	58.6	1.9	1000	US-10-021-698A-291	Sequence 291, App
10	55.6	1.8	536	US-10-779-543-17721	Sequence 17721, A
11	55	1.7	27890	US-60-548-091-5662	Sequence 5662, Ap
12	55	1.7	126872	US-60-548-091-5703	Sequence 5703, Ap
13	54.4	1.7	1035	US-10-779-543-8023	Sequence 8023, Ap
14	50.8	1.6	2637	US-10-724-972A-995	Sequence 995, App
15	50.8	1.6	327636	US-60-550-051-21294	Sequence 21294, Ap
16	50.4	1.6	201	US-60-550-051-21273	Sequence 21273, A
17	50.4	1.6	13485	US-60-550-051-3088	Sequence 3088, Ap
18	49.8	1.6	2100	US-10-416-330-17	Sequence 17, Appl
19	49.6	1.6	1450	US-10-779-543-7878	Sequence 7878, Ap
20	49.2	1.6	5917	US-10-808-727-9	Sequence 9, Appl
21	48.8	1.5	30612	US-10-724-972A-326	Sequence 326, App
22	48.6	1.5	201	US-60-550-051-20553	Sequence 20553, A
23	48.6	1.5	27890	US-60-548-091-5662	Sequence 5662, Ap
24	48.6	1.5	126872	US-60-548-091-5703	Sequence 5703, Ap
25	48.4	1.5	2226	PCT-US04-05654-1714	Sequence 1714, Ap
26	48	1.5	500	PCT-US04-02000-562	Sequence 562, App

27 47.4 1.5 2564 6 US-10-416-330-12
28 47.4 1.5 170546 6 US-10-767-471-10703
29 47.4 1.5 41665 6 US-10-767-471-10744
30 46.6 1.5 1035 6 US-10-779-543-4437
31 46.6 1.5 1102 6 US-10-767-795-3060
32 46.4 1.5 1423 6 US-10-767-795-3350
33 46.4 1.5 3053 6 US-10-491-213-108
34 46.2 1.5 760 6 US-10-021-698A-161
35 46.2 1.5 780 6 US-10-021-698A-540
36 46 1.5 885 6 US-10-767-701-20477
37 46 1.5 40000 6 US-10-796-280-12600
38 46 1.5 107450 6 US-10-796-307-8822
39 45.8 1.4 173995 6 US-10-796-307-8746
40 45.6 1.4 3636 6 US-10-724-972A-1890
41 45.6 1.4 107191 7 US-60-550-051-2971
42 45.4 1.4 198285 6 US-10-775-163-338
43 45.2 1.4 201 7 US-60-550-051-20567
44 45.2 1.4 3885 6 US-10-377-636-3
45 45 1.4 481 6 US-10-767-701-17770

ALIGNMENTS

RESULT 1

US-10-486-020-51
; Sequence 51, Application US/10486020
; GENERAL INFORMATION:
; APPLICANT: AZIMZAI, Valda; BARROSO, Ines;
; APPLICANT: BAUGHN, Marian R.; BECHA, Shanya D.;
; APPLICANT: BORSKY, Mark L.; DUGGAN, Brendan M.;
; APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.;
; APPLICANT: FORTY, Ian J.; GIETZEN, Kimberly J.;
; APPLICANT: GORVAD, Ann E.; GRAUL, Richard C.;
; APPLICANT: GRIFFIN, Jennifer A.; GURURAJAN, Rajagopal;
; APPLICANT: HARELIA, April J.A.; ISON, Craig H.;
; APPLICANT: KABLE, Amy E.; KHAN, Farrah A.;
; APPLICANT: LEE, Sally; LEE, Soo Yeun;
; APPLICANT: LI, Joana X.; REDDY, Roopa;
; APPLICANT: RICHARDSON, Thomas W.; SPRAGUE, William W.;
; APPLICANT: SWARNAKAR, Anita; TANG, Y. Tom;
; APPLICANT: WARREN, Bridget A.; XU, Yuming;
; APPLICANT: YAO, Monique G.; YUE, Henry;
; APPLICANT: YUE, Huibin
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PF-1126 USN
; CURRENT APPLICATION NUMBER: US/10/486,020
; PRIOR APPLICATION NUMBER: PCT/US02/25465
; PRIOR FILING DATE: 2004-02-05
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/311,017
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/313,070
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/313,071
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/314,678
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/316,692
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/317,913
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/322,182
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: US 60/340,747
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/342,761
; PRIOR FILING DATE: 2001-12-20
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PERL Program
; SEQ ID NO 51
; LENGTH: 3108

Sequence 12, Appl
Sequence 10703, A
Sequence 10744, A
Sequence 4437, Ap
Sequence 3060, Ap
Sequence 3350, Ap
Sequence 108, App
Sequence 161, App
Sequence 540, App
Sequence 20477, A
Sequence 12600, A
Sequence 8822, Ap
Sequence 8746, Ap
Sequence 1890, Ap
Sequence 2971, Ap
Sequence 338, App
Sequence 20567, A
Sequence 3, Appl
Sequence 17770, A

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7506167CB1
US-10-486-020-51

Query Match      91.0%; Score 2880.4; DB 6; Length 3108;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 3023; Conservative      0; Mismatches      6; Indels 129; Gaps 1;

QY 1 AGGTTTATCCATGAAGAGACAGCAATTTAAAGTGTTATTCACCAAAAAATAAGCTT 60
DB 72 AGGTTTATCCATGAAGAGACAGCAATTTAAAGTGTTATTCACCAAAAAATAAGCTT 131
QY 61 CAAATATGTGATGTGAAACTGCCAGAACTAAGGCGCGCGCGCTCAGACCAGCGCTG 120
DB 132 CAAATATGTGATGTGAAACTGCCAGAACTAAGGCGCGCGCGCTCAGACCAGCGCTG 191
QY 121 CTCAGATGTAAAGTGTACAGAGAGCGCCAGGGAGGTGGTGGGGAACAATGGGCTG 180
DB 192 CTCAGATGTAAAGTGTAAAGAGAGCGCCAGGGAGGTGGTGGGGAACAATGGGCTG 251
QY 181 TGAGGCGCTGGGCTGCGCGGTTCCCGAGCTCCCGCCCGAGCCCGCTCCACAGTGGTGG 240
DB 252 TGAGGCGCTGGGCTGCGCGGTTCCCGAGCTCCCGCCCGAGCCCGCTCCACAGTGGTGG 311
QY 241 CTCGGTGGTGGTGCACGTCGCGATTCGGGTTCAGACCAGCGCTGCGTTCCTCCACC 300
DB 312 CTCGGTGGTGGTGCACGTCGCGATTCGGGTTCAGACCAGCGCTGCGTTCCTCCACC 371
QY 301 GCTTGTGTGGCCAGTGTACTCGGTGACCGCCAGAGCAGCTGACGCTATGAGGAG 360
DB 372 GCTTGTGTGGCCAGTGTACTCGGTGACCGCCAGAGCAGCTGACGCTATGAGGAG 431
QY 361 CTTGTGTACACCTCAGCCCTACCTGGGCTGGTCTGGAGAGCTACCGAGAGTTGTG 420
DB 432 CTTGTGTACACCTCAGCCCTACCTGGGCTGGTCTGGAGAGCTACCGAGAGTTGTG 491
QY 421 GCAGCACTACCTGAGAGTATGACACAGATGAGATCCTTATGTTTTCATCGGAACTG 480
DB 492 GCAGCACTACCTGAGAGTATGACACAGATGAGATCCTTATGTTTTCATCGGAACTG 551
QY 481 GTGGTATGTGACCTGTATTTGTTGTTTCTCTCTTTTGTGGAGAGTTT 540
DB 552 GTGGTATGTGACCTGTATTTGTTGTTTCTCTCTTTTGTGGAGAGTTT 611
QY 541 AGATCGTTAGGAGTGGCTTTACGTGGAGAGAGCAAACTTGGTCCACGCTTCT 600
DB 612 AGATCGTTAGGAGTGGCTTTACGTGGAGAGAGCAAACTTGGTCCACGCTTCT 671
QY 601 GGAATAATTGAAGAAAATGTAACTACTTGAAAAATTTAGCCTTATTCAAAAAGATAT 660
DB 672 GGAATAATTGAAGAAAATGTAACTACTTGAAAAATTTAGCCTTATTCAAAAAGATAT 731
QY 661 GAGGCTATGAGTATGATCATCTTTAGAGGATGCCAGCTTTGAGAGAGGAGCAGAGAA 720
DB 732 GAGGCTATGAGTATGATCATCTTTAGAGGATGCCAGCTTTGAGAGAGGAGCAGAGAA 791
QY 721 GAAGCAGAGTGTGGAGCAACCTGTGAAAAAGCTGAAACAGGTCCCAATTTCTGAACCTTG 780
DB 792 GAAGCAGAGTGTGGAGCAACCTGTGAAAAAGCTGAAACAGGTCCCAATTTCTGAACCTTG 851
QY 781 GATGAATCTCTGTCTAGAAAAGACTTTAAACCAAGAGAAATCTTAAACATTTCTCAACAA 840
DB 852 GATGAATCTCTGTCTAGAAAAGACTTTAAACCAAGAGAAATCTTAAACATTTCTCAACAA 911
QY 841 GATGAATGTGCGGATATTTCAAAAAGTATACAGTCTCTAGAGATGAGTCAAAATCC 900
DB 912 GATGAATGTGCGGATATTTCAAAAAGTATACAGTCTCTAGAGATGAGTCAAAATCC 971
QY 901 CTCAATCAAAATAGCTGAAGCCAAATCATCTGCAAGACATTTAAATGATGTAAGAA 960

```

```

DB 972 CTCAAATCAAAATAGCTGAAGCCAAATCATCTGCAAGACATTTTAAATGAGTGAAGAA 1031
QY 961 CGACGGCTATAGCAATTAAGAGATGCTTTGAAATGAAAAATTTCTCAACTTCAGACAAGCCAT 1020
DB 1032 CGACGGCTATAGCAATTAAGAGATGCTTTGAAATGAAAAATTTCTCAACTTCAGACAAGCCAT 1091
QY 1021 AAAACAGCTTTTTCAGCAAGAAAGCTGAAGTATGAAAGGAGAAAGTGAAGTGAACCTTAATAAA 1080
DB 1092 AAACAGCTTTTTCAGCAAGAAAGCTGAAGTATGAAAGGAGAAAGTGAAGTGAACCTTAATAAA 1151
QY 1081 CAGAAAATAACATTTGAAGACTCCAAAGTACACGCGAACAAGTTCTGAATGATAAGAA 1140
DB 1152 CAGAAAATAACATTTGAAGACTCCAAAGTACACGCGAACAAGTTCTGAATGATAAGAA 1211
QY 1141 AATCACATCAAGACCTTGAAGTATGCACTTGCCTAATGATGAAAGATCAGGCTGCTGCTT 1200
DB 1212 AATCACATCAAGACCTTGAAGTATGCACTTGCCTAATGATGAAAGATCAGGCTGCTGCTT 1271
QY 1201 GAAGAAACACACACCGATGATGATTAACCTGGAATTAGAAGTGAACAGTGAATCGGAAAT 1260
DB 1272 GAAGAAACACACACCGATGATGATTAACCTGGAATTAGAAGTGAACAGTGAATCGGAAAT 1331
QY 1261 GGTCTTACTTAGATAATCTCCAAAGAGAGCTTTGAAGAACTGATTTCATGCTGCTAAG 1320
DB 1332 GGTCTTACTTAGATAATCTCCAAAGAGAGCTTTGAAGAACTGATTTCATGCTGCTAAG 1391
QY 1321 TTAATGCTTTTAAACCTTAGAAGGAGAAAGAAACCAATTTATATTCAGTTGCTT 1380
DB 1392 TTAATGCTTTTAAACCTTAGAAGGAGAAAGAAACCAATTTATATTCAGTTGCTT 1451
QY 1381 GAAGTTCATTAACCAAGAGAGAGCTTACAGAGCATTTAAATAATCTTCAGACTCAACAA 1440
DB 1452 GAAGTTCATTAACCAAGAGAGAGCTTACAGAGCATTTAAATAATCTTCAGACTCAACAA 1511
QY 1441 GCATCTTTGAGTCAGAAAAACACATTTTGAATAAGAGAAATCAGAACTTCAACAGAAA 1500
DB 1512 GCATCTTTGAGTCAGAAAAACACATTTTGAATAAGAGAAATCAGAACTTCAACAGAAA 1571
QY 1501 CTTAAAGTAAATGCTGAATTTATCAAGAAAATGAATAAATCTCCACAGAAATTAACA 1560
DB 1572 CTTAAAGTAAATGCTGAATTTATCAAGAAAATGAATAAATCTCCACAGAAATTAACA 1631
QY 1561 GTAGAGAAAATTTATCGGTTAGAGAAAAGAGAAACTTTCTAAAGTGAATGAAAAGATC 1620
DB 1632 GTAGAGAAAATTTATCGGTTAGAGAAAAGAGAAACTTTCTAAAGTGAATGAAAAGATC 1691
QY 1621 AGCCATGCCACTGAGAGCTGGAGCCTATAGAAAGCGAGCCAAAGATCTTGAGAGAA 1680
DB 1692 AGCCATGCCACTGAGAGCTGGAGCCTATAGAAAGCGAGCCAAAGATCTTGAGAGAA 1751
QY 1681 TTGAGAGAACTATTCTTATCAAGGCGAGATTTTCCCATGAGAAAAAGCAGAT 1740
DB 1752 TTGAGAGAACTATTCTTATCAAGGCGAGATTTTCCCATGAGAAAAAGCAGAT 1811
QY 1741 GATAAATTTGGTGGCAGCTCGGAATGCTGAAGAAACCTCAATGTTTAAAGAAAGAAAT 1800
DB 1812 GATAAATTTGGTGGCAGCTCGGAATGCTGAAGAAACCTCAATGTTTAAAGAAAGAAAT 1871
QY 1801 GCTCACACAGACAAAAATTAACCTGAAACAGAGCTTAAATTTGAACCTTTTGAAGAAAGAT 1860
DB 1872 GCTCACACAGACAAAAATTAACCTGAAACAGAGCTTAAATTTGAACCTTTTGAAGAAAGAT 1931
QY 1861 CTTTATGCACTCGATGTTCCAAATACAGCATTTGGCAGAGAGCATTTCCCATATGTTGCC 1920
DB 1932 CTTTATGCACTCGATGTTCCAAATACAGCATTTGGCAGAGAGCATTTCCCATATGTTGCC 1967
QY 1921 TCACCAATTTGGGTTGGCCTTCTATCTGAAAAAAGAGCTTTTCTCTCTCTCTCTCTCTCT 1980
DB 1968 ----- 1967
QY 1961 GAGGCTCACTCAGACTCTCACCTTTGCTTCCAGGGGAGAGAGAGAGGCTCAGAGGC 2040
DB 1968 -----AGAGGCTCAGAGGC

```


QY 210 CTCCCGGCGAGCGCTCCACAGTGGTCCGCTCCGGTTGGTTGTCACTGCGCATTCGG 269
Db 149 TCCCGGCGAGCGGCTCCGAGTGGTCCACTCCGGT---TGCGGGGTGGGATTCGG 204
QY 270 GTTCCAGACCCAGGCTGGGTCTTCCACCGGTTCTTGT-----GCCAGTGT 319
Db 205 GTTCCGAGCCGAGGCTGTGTCTTCCCGCGTTTATTTGGCGCCCGACAGCGCGGGTT 264
QY 320 ACTGCGGTACCGCCAGAGAGAGCGCTCGAGCTATGGAGAGCGTGGTCTACCCCTCAGC 379
Db 265 ACTGTGGCGACACGAGAGAGAGCTTTGGCGCTATGGAGAGCGCCGGGGTACCCCTCAAC 324
QY 380 CTTACCTGGGCGTGGTCTTGGAGAGCTACGAGAGTTGTGCGAGCACTACCTGAGAGTA 439
Db 325 CGTATTGGGCGTGGTCTTGGAGAGCTACCGAGGTTGTGCGAGCACTGCCCTGAAGTA 384
QY 440 TGAGACAGATGAATCTTATGTTTCCATCGGAAGTGGTGGTATGCGAGCTGTA 499
Db 385 TGAGACAGATCTTATCTTTATGTTTCCATGGGAATGGGATATGCGAGCTGTTG 444
QY 500 TTGGATTTTGTGTTCTCTTTTGTGGAGAGTTTATAGTCGGTTAGAGTCGCG 559
Db 445 TTGGATTTTGTGTTCTCTTTTGTGGAGAGTTTATAGTCGGTTAGAGTCGCG 504
QY 560 TTTACGTGGAGAGAGCAAACTTGGTGCACGCTTCTGACCTAAATGGAGAAAT 619
Db 505 TTTATGTGGAGAGAGAAAGCTTGTCTTATGCTTTCTGACTTAATGAAGAAAAA 564
QY 620 GTAACTACTTGAATAATTTAGCTTTATCAAAAGAGTATGAAGCTATGAAGTAGCT 679
Db 565 GTAACTACTTGAATAATTTAGCTTTTCAAAAGAGTATGAAGCTATGAAGTAGCT 624
QY 680 CATCTTTAGAGATGCCAGCTTTGAGAGAGGAGCAGAGAGCAAGAGTTGGAG 739
Db 625 CATCTTTAAAGAGTGCAGCTTTGAGAGAGGAGCAAC--AGAGAGCAAGAGTTGGAG 681
QY 740 CAACCTGTGAAGAGCTGAAAGAGTCTGAATCTGAAGTGAATCTCTGTCTAG 799
Db 682 CAACCTGTGAAGAGCTGAAAGAGTCTGAATCTGAAGTGAATCTCTGTCTAG 741
QY 800 AAAAGACTTAAACAGAGAAATCTAAACATTTCTCAAGATGAATGATGGCGGATA 859
Db 742 AAAAGAGTAAAGAGAGAGAAATCCAAATCTGAACAGATGAATGATGGCGGATA 801
QY 860 TTTCAAAAGTATACAGTCTCTAGAGATGAGTCAAAATCCCTCAATCACAATAGCTG 919
Db 802 TTTCAAAAGATACAGTCTCTAGAGATGAGTCAAAATCCCTCAATCACAATAGCTG 861
QY 920 AAGCAAAATCATCTGCAAGACATTTAAATGAGTGAAGAAACGAGGCTATAGCAATAA 979
Db 862 AAGCAAAATGACCTTCAAGATATTTCAAAATGAATGAAGAAACGACTGAAGTAGCAATAA 921
QY 980 AAGATGCTTTGAATGAATAATCTCACTTCAGACAGCCATAACAGCTTTTTCAGCAAG 1039
Db 922 AAGATGCTTTGAATGAATAATCTCACTTCAGAGAAAGCCAGAACAGCTTT---GCAAG 978
QY 1040 AAGTGAAGTATGAAGAGAGAGTGAAGTGAATTAATAACAGAGAAATAACATTTGAAG 1099
Db 978 AAGTGAAGTATGAAGAGAGAGTGAAGTGAATTAATAACAGAGAAATAACATTTGAAG 1038
QY 1100 ACTCAAGTACCGAGACAGTCTGATGATGAAGAAATCAATCAGAGCCCTGA 1159
Db 1039 ACTCAAGTATGAGAGAGAGTCTGATGATGAAGAAATCAATCAGAGCTCTGA 1098
QY 1160 CTGACACTTCCCAATGATGAAGATCAGGCTGTGCTTGAAGAGAGACACACGAGT 1219
Db 1099 CTGAACGCTTGTAAAGATGAAGATTTGGGCTGTATGCTTGGAGAGAGACATAACGAGT 1158
QY 1220 ATGATACCTGGAATTAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1279
Db 1159 ATGATACCTGGAATTAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1218
QY 1280 CTCAAAAGAGAGCTTTGAAGAAAGTGAATTCATGCTGCTGAAGTAAATGCTTTAAAAA 1339

Db 1219 CTCAAAAGAGAGCTTTGAAGAAAGTGAATTCATGCTGCTAGTTAAATGCTTTAAAAA 1278
QY 1340 CTTTAG 1399
Db 1279 CTTTAG 1338
QY 1400 AAGAGCTTACAG 1459
Db 1339 AAGAGCTTACAG 1398
QY 1460 ACACACATTTTGAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1519
Db 1399 ACACACATTTTGAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1458
QY 1520 TATATCAAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 1579
Db 1459 TATATCAAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 1518
QY 1580 TAGAGAAAG 1639
Db 1519 TAGAGAAAG 1578
QY 1640 TGGAGAGCTATAG 1699
Db 1579 TGGAGAGCTATAG 1638
QY 1700 CTTATCAAGGGAG 1759
Db 1639 CTTATCAAGGGAG 1698
QY 1760 GGAATGCTGAAG 1819
Db 1699 GGAATGCTGAAG 1758
QY 1820 TAACTGAAG 1879
Db 1759 TAACTGAAG 1818
QY 1880 CAATATAG 1939
Db 1819 CAATATAG 1878
QY 1940 CATCTGAAG 1999
Db 1879 CATCTGAAG 1938
QY 2000 CACCTTTGCTTCCAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2059
Db 1939 CACCTTTGCTTCCAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1998
QY 2060 ATCAGATTAACCAATGAAG 2119
Db 1999 ATCAGATTAACCAATGAAG 2058
QY 2120 CTCCTCTGACACTGGGCTCTGCTCACTCATGGAGAGAGAGAGAGAGAGAGAGAGAGAG 2179
Db 2059 CTCCTCTGACACTGGGCTCTGCTCACTCATGGAGAGAGAGAGAGAGAGAGAGAGAGAG 2118
QY 2180 CTCGCCAGGAG 2239
Db 2119 CTCGCCAGGAG 2178
QY 2240 GTTCTAAATCTGGTAG 2299
Db 2179 GTTCTAAATCTGGTAG 2238
QY 2300 TGGATAAATGAGTGGGTCAATGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2359
Db 2239 TGGATAAATGAGTGGGTCAATGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2298
QY 2360 ATGATCTTGGTAAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2419